

Dangers of over-dependence on peer-reviewed publication

A number of independent trends are increasing the significance of journals' roles in providing high-quality information. Other sources need to be strengthened.

Like it or not, the days of ink on paper are numbered — and the number is smaller than many people imagine. This may sound regrettable: no one can comfortably read even a laptop computer in bed, whereas *Nature's* Millennium Essays, for example, are just the sort of erudite relief that hard-driven scientists deserve as they drift towards sleep. But lightweight, flexible screens with print-like quality, instant page upload, high-speed connectivity and large memories will arrive, and such nocturnal browsing could then include the opportunity, with this week's essay (see page 745), to hop directly to digitized versions of the works of Henri Poincaré and others who so variously interpreted Heinrich Hertz's sparks. If the benefits were for scientists alone, that technology would take decades to arrive. The fact that everybody will want it explains why Bill Gates is spending so much of his time with telecommunications companies.

Preprint servers have prompted suggestions that electronic networks could liberate researchers from the stranglehold of publishers. Yet the dependence on established gatekeepers is increasing. Electronic communications are just one of several reasons.

As access to scientific information becomes either free or readily purchasable per view, rather than by annual subscription, the Internet will make primary scientific information easily available not only to scientists but also to the many other stakeholders in science. That is all to the good given the intrinsic desirability of openness as well as the increasingly accepted need to involve (rather than merely inform) various publics in technology regulation and risk assessment.

Dangers of the Internet

But the risks with such access are clear too. The results of Arpad Pusztai published by *The Lancet* last week (see page 731) provide a timely example. Given that Pusztai's initial claims on a television programme were welcomed and exploited by lobby groups, and that the Royal Society felt it necessary to set up a committee to investigate them in the absence of a peer-reviewed publication, one can only assume that such claims announced on a scientific preprint server would have been equally open to burdensome contention. Preprints are ripe for misappropriation in such public controversies. The dangers with biomedical research are also obvious, as is the risk of exploitation of a preprint server to establish premature claims for priority, whether for research competitiveness or commercial interests.

With such concerns in mind, the announcement last week of an agreement between the proposed electronic publication server PubMed Central and the US National Academy of Sciences has the merit of distancing itself from preprints (see page 733). In contrast, there was a joint proposal in Frankfurt last week by the American Association for the Advancement of Science and the International Council of Scientific Unions to the Association of Science, Technical and Medical publishers that should set off alarm bells. Responding to developments in electronic networks, it proposes the recognition of two stages of publication: "first publication" — exposure in some permanent form allowing an establishment of priority — and the

"definitive publication" — a refereed version. In the light of the above concerns, that blurring of the concept of "publication" is exactly what the world does not need.

Media reliance

The mass media provide another example — and a worrying one — of the increasing significance of peer-reviewed publication. Their coverage of research provides no guarantee of the science's quality but is nevertheless much valued by scientifically fascinated readers, listeners and viewers, as well as by the scientists whose work is highlighted and (Pusztai apart) their employers. Accordingly, journalists increasingly find themselves supplied with press releases by journals. But, especially with the pressures associated with the mass media's own electronic networks, they have less and less time to research the stories in depth. In television, above all, the drive to cut costs in filling air time is ever more intense, and can only undermine the quality of broadcast coverage. The implication is that scientific journals should supply journalists not only with a balanced lead to every story but with details of its context too, comprehensively presented. That is a tall order.

Dependence on the journals, especially those with high impact factors, has seemingly never been greater within the scientific community itself, in the assessment of staff and their institutions. In some countries this is taken to extremes: in order to encourage an outward-looking approach, some universities in China and Taiwan pay their staff bonuses following publication in journals tracked by the Institute for Scientific Information — the more important the journals, the higher the bonus, in some cases. Government grants can automatically follow publication in the top journals. In the United States, the award of tenure can be strongly influenced by such publication. Despite the avoidance of crude paper counting in quality assessment, the pressure on assessors' time as well as that of institutions and individual academics means that publication in a prestigious journal becomes ever more convenient as a shorthand indicator of achievement.

Peer-reviewed publications, as everybody in the business knows, include papers that are definitive and papers that are highly controversial. The editor of *The Lancet* has attracted vehement criticism not only because Pusztai's work has demonstrable technical shortcomings and thus falls below the normally high standards that he maintains, but also because he is perceived to be undermining the growing dependence on journals like his in an increasingly controversial arena.

Scientists and their institutions are in for a more turbulent future as access to their information becomes ever wider. The journals should be expected to maintain their standards in publishing valid, if occasionally credibility-stretching, science. But the ever-increasing reliance on them for quality control has disadvantages that should be countered by adequate provision of time and resources for independent assessment and, in the midst of controversies, publicly funded agencies providing comprehensible, reliable and prompt complementary information over the networks.





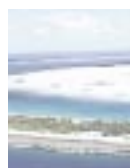
Universities

Canada is to create 2,000 research chairs, says prime minister
p731



Space

NASA and European Space Agency plan science missions
p734



Nuclear tests

US Senators ignore scientific advice on test ban treaty
p735



Neptune

Earth-based cameras put planet in sharp focus
p737

Venter's *Drosophila* 'success' set to boost human genome efforts

Paris

Maverick, dreamer or visionary? The jury is still out on Craig Venter, head of Celera Genomics Systems, whose ambition is to sequence and piece together the entire human genome by 2001 using a controversial 'whole-genome shotgun' strategy.

But Venter, who thrives on his image as the *enfant terrible* of genomics, seems poised to score a victory over sceptics who predict that shotgun sequencing will not work with large, complex genomes.

Last month, Celera finished sequencing the 180-million-base-pair (Mbp) genome of *Drosophila melanogaster*. The company intends to publish the complete genome at the beginning of 2000, in collaboration with the publicly funded *Drosophila* Genome Project, headed by Gerry Rubin at the University of California at Berkeley.

Does this mean that Venter will repeat the feat in humans? Succeeding with *Drosophila* would be a convincing proof of the feasibility of shotgunning large genomes, but the ultimate test will be the 3,500-Mbp human genome. But Celera is unlikely to sequence the human genome using only a shotgun strategy, as it will almost certainly be quicker to combine this data with conventional maps generated by the publicly funded international Human Genome Project (HGP).

Indeed, if there is one lesson from the *Drosophila* project it is that the collaboration between Celera and Berkeley has speeded up the assembly of the genome and greatly improved the quality of the data obtained.

The Berkeley team have produced 26.5 Mbp of sequence and a low-resolution physical map that will help Celera align its shotgun sequences along the genome. Celera has provided a '10×' sequence of the entire genome, which means that ten bases of sequence have been generated for every base of genome. The completeness of a genome increases with the number of times it is sequenced. The norm is 5–6×, and 10× is the current gold standard. One scientist who has seen Celera's unpublished 10× *Drosophila* data says it is "terrific".

Ironically, in the race to complete the

human genome, Francis Collins, director of the US human genome effort, and Michael Morgan, head of genomics at the Wellcome Trust in the United Kingdom, recently revised their goal from a 10× map by 2003 to a 5× draft by spring 2000.

The *Drosophila* sequence is quite an achievement. "The *Drosophila* community is very grateful that this speed has been possible, and the input of Celera made a big difference," says Allan Spradling, an embryologist at the the Carnegie Institution of Washington and the Howard Hughes Medical Institute.

Shattering a genome into fragments for sequencing is easy. Arranging the fragments in the order in which they occur on the chromosome is more difficult. Celera's big achievement is that it seems to be making solid progress in piecing together the 3.2 million sequence fragments of the fruitfly genome.

In the conventional 'clone-by-clone' approach being taken by the HGP, each fragment is cloned and propagated in a library by inserting it in the genome of a bacterial artificial chromosome (BAC). The jumbled

Wellcome funds cancer database

London

Britain's Wellcome Trust is to provide up to £10 million (\$16.5 million) over the next five years to detect the genes that are mutated in human cancers and release the information to a public database.

The project, to be based at the Sanger Centre near Cambridge, will use sequence information from the international Human Genome Project — of which the trust is providing one third of the cost — as the basis for screening genomes for the mutations underlying oncogenesis.

The work is "one of the first 'next step', post-genome projects," says Mike Dexter, director of the trust. "Although it will focus on cancer-related genes, it will also act as a 'proof of concept' for ideas that can be applied to most, if not all, polygenic diseases."

The idea of screening cancer cell genomes for mutations was proposed by Mike Stratton and Richard Wooster of the Institute for Cancer Research (ICR) in Sutton. They led the team that located and then identified the *BRCA2* breast-cancer susceptibility gene, based on a sequence



Wooster and Stratton: looking for mutations.

from the Sanger Centre.

"The completion of the Human Genome Project will enable us to develop technology that we can use to compare the DNA sequence of normal and cancerous tissue from the same individual and hence

identify the mutated genes implicated in tumour development," says Stratton.

Providing substantial support for cancer research is a new move for the Wellcome Trust. But Dexter argues that the basic nature of the work, its direct link to the sequencing project, and its relevance to other polygenic diseases make it suitable for the trust's support. Additional funding (£1 million initially) is being provided from the ICR, and further funding is being sought.

Data will be made accessible to all researchers, in line with trust policy. "Initially we will be focusing on the common cancers, but ultimately the aim is to provide as much data as possible about the broad spectrum of human neoplastic disease," says Stratton, who will lead the project at the Sanger Centre.

David Dickson

set of chunks, or 'contigs', are then rearranged into a physical map, typically by looking for overlapping fragments sharing short sequences of DNA.

Once BAC contigs spanning the genome have been constructed, the sequences of the 40- to 400-kilobase-pair (kbp) BACs are compiled, yielding the sequence of each chromosome and eventually the entire genome. The HGP requires building a library of some 30,000 BACs.

Venter is taking a different tack, processing the millions of DNA fragments, or 'reads', directly and feeding them into a supercomputer. Gene Myers, head of Celera's computational biology group, is developing software to assemble the jigsaw by matching their sequences.

Until 1995, the upper limit for such whole-genome shotgunning was 40 kbp. Then Graham Sutton at Venter's Institute for Genomic Research in Rockville, Maryland used the technique to sequence the 1,800-kbp genome of the bacterium *Haemophilus influenzae*. Celera now appear to have pushed the upper limit through the ceiling.

Myers presented data at the eleventh International Genome Sequencing and Analysis Conference in Miami, Florida last month. He claimed that, in a test run on a 6 × 28-Mbp stretch of *Drosophila* sequence prepared at Berkeley using a clone-by-clone approach, a scaffold of contigs generated by his computer assembler gave a perfect match. The only discrepancy was traced back to an erroneous chimaeric clone in the Berkeley data set.

The demonstration has won over some critics of whole-genome shotgunning. "The tests they did were stringent. It is hard to argue with a match like that," says Spradling. "It is impressive." Michael Ashburner, from the European Bioinformatics Institute in Cambridge, says: "I'm now fairly convinced the technique works."

Difference in complexity

But others remain unconvinced. Philip Green, a biocomputing expert at the University of Washington, says: "If they get it right in *Drosophila* I'll be impressed, but it will not persuade me that they will succeed in humans." This view is shared by both John Sulston, head of the Sanger Centre, in Cambridge, and Francis Collins, director of the US National Human Genome Research Institute.

They point out that the human genome is larger and more complex than that of *Drosophila*. Green says the human genome contains millions of repeat sequences. He is sceptical that Myers will be able to position these, arguing that doing so is easier using a clone-by-clone approach, where each clone will have fewer copies of any given repeat.

But Myers insists that the *Drosophila* demonstration "proves that we are not being



Myers: claims his software gives a perfect match, spots repeats and creates a virtual genome.

confounded by the repeats. We are able to identify all the unique stretches of the genome, assemble them and order them without any mistakes."

Myers claims that his software is able to spot repeats in the genome. He also says he has created virtual genomes on screen, incorporating all sorts of nasty repeats, and that his assembler is able to deal with them.

Myers admits, however, that his software — which runs to 150,000 lines of code — is incomplete. He is confident that he can finish the software within weeks.

Much of his confidence stems from the decision by Celera to use a variant of classical shotgunning, called double-barrelled shotgunning. Most sequencing projects collect data from only one end of a DNA fragment but Celera is using sequence from both ends.

If you take a 2,500-bp insert and read 500 bp off each end, you know that you have a 'mated' pair of reads which, in the final assembly, should be roughly 1,500 bp apart and pointing at each other. This is a powerful tool, which in principle allows assembly across repeats.

Myer says: "Normally if I shotgun *Drosophila* I get 3.2 million reads. Period. You know nothing about them. What we get using mates is 3.2 million reads where for a great number they come in pairs which are at a certain distance apart in the genome. That is a significant additional clue about how to do the assembly."

Mates are commonly used to check genome assemblies, but have not been widely used as core assembly strategies because they require a quality of sequence data that has been difficult to achieve using conventional slab gel machines, which typically give ten per cent false pairing of mates. Celera has invested in some 250 capillary electrophoresis systems — Perkin Elmer's ABI Prism 310 Genetic Analyzer — which cut the error rate to between 1 and 0.1 per cent.

But Green reckons that Myers will not be able to solve the problem of large, low-frequency repeats. Ironically, the easiest way for Celera to resolve large repeats will be to use long-range BAC data generated by the HGP. Myers admits that he will probably need the 600,000 BAC end sequences from the HGP to detect these longer repeats. Indeed, to speed things up, Venter will not rely solely on whole-genome shotgunning, but will also use the mapping data flowing from the HGP. Green predicts that Venter will map the HGP reads to the rough BAC contig map.

Collaboration

Many genome scientists believe that the HGP should follow Rubin's lead and strike a deal with Celera. But when Myers and James Weber, director of medical genetics at the Marshfield Medical Research Foundation in Wisconsin, first proposed a whole-genome shotgun of the human genome in 1997, it was perceived as an alternative, and a threat, to the public clone-by-clone approach.

This has left smouldering resistance to Venter, says Sulston, who argues that the two approaches are complementary. "Myers is doing the human genome in a larger and bolder way and it is excellent that this experiment is being done; but we need both approaches."

Sulston says: "By continuing the robust HGP in the public domain using a clone-by-clone approach we ensure that the thing is produced in a timely and complete way. It is an insurance policy against failure at Celera, and to ensure that the data is made public."

"If Celera combined their data publicly with us that would be the best possible option." But attempts at formal collaboration have stalled, mainly because of the HGP's demand that Venter abide by its rule to release data within 24 hours (see *Nature* **397**, 93; 1999). "We would like full collaboration. We cannot because the sequence must be fully and openly released," says Sulston.

Paul Gilman, head of policy planning at Celera says that "Celera continues to believe that through cooperation we could accelerate the completion of the genome, not a draft, to even earlier than Celera's 2001 completion date."

Venter shipped the *Drosophila* sequence to commercial subscribers to his database on 4 October. The deal with Rubin provides for the data to be submitted to Genbank, perhaps as soon as this month. The annotation of the *Drosophila* genome will also begin next month, when 20 Celera scientists and 30 scientists from the *Drosophila* community meet for an "annotation jamboree".

Rubin says that the annotation data will now also be submitted to the Genbank database on publication of the genome in early 2000. "The entire annotation record will be provided to FlyBase for public release without restriction."

Declan Butler

Journal under attack over controversial paper on GM food

London

The medical journal *The Lancet* has been criticized by Britain's Royal Society for publishing a paper on genetically modified (GM) potatoes by controversial author Arpad Pusztai.

The study has been a source of continuing furore in Britain following a statement by Pusztai to the press last year — before any findings had been published that GM foods may stunt the growth of rats. This statement triggered widespread concern over the issue of GM foods. The row continued when the Royal Society reviewed Pusztai's data and concluded that the study was based on flawed design, execution and analysis (*Nature* 398, 98 & 399, 188; 1999).

The Lancet's editor, Richard Horton, says the journal raised the threshold for publication of this paper by sending it to six reviewers instead of three, the number normally used. This was because "Pusztai had recklessly made claims about his data a year ago and because of the issue's sensitivity". The paper is accompanied by a commentary by researchers at Wageningen Agricultural University in The Netherlands who are critical of the study's design and conclusions.

But Horton defends the journal's decision to publish the paper, saying that it was made on scientific grounds. "A majority supported publication, but for different reasons. If we hadn't published, it wouldn't be surprising if someone accused us of censorship."

Horton says that one of the benefits of publication is that Pusztai "has had to retract his original claim because his data, which we've published, absolutely don't show that genetically modified foods stunt the growth of rats".

Aaron Klug, president of the Royal Society, said the society would not have published the paper because "it confirms the society's judgement that the experiments were flawed". He said publication had given the study undeserved authenticity.

Peter Collins, the society's senior policy adviser, accuses *The Lancet* of taking "a very strong position on this. It is clear from an editorial previously published by *The Lancet* that the editor or the magazine were strongly supportive of Pusztai before they had seen his science. I don't think *The Lancet* can present its decision to publish Pusztai's paper now as a routine scientific decision." **Natasha Loder**

Japanese guidelines specify the terms of gene patents

Tokyo

The Japanese Patent Office (JPO) has released guidelines on human gene patents. These outline its view of the patentability of expressed sequenced tags (ESTs), single nucleotide polymorphisms (SNPs) and full-length complementary DNA (cDNA) clones.

Although the office reached a general agreement with the US Patent and Trademark Office and the European Patent Office over EST patents in July — namely that such partial DNA sequences cannot be patented without demonstrated utility — its position on other genetic patents had remained unclear.

The new guidelines, unveiled on 5 October, are a response to a call for guidance from Japanese researchers and industry, particularly in the wake of the government's decision to increase support for biotechnology (see *Nature* 400, 389; 1999).

This initiative includes a planned database of SNPs in the Japanese population, which is hoped to lead to novel drugs and diagnostic techniques, and a project aimed at creating a central repository of cDNA clones for medical and research applications.

The introduction of a technology-transfer bill, based on the US Bayh-Dole legislation, which grants private patents on government-funded research, is also hoped to encourage patent applications from indus-

try-university collaborations. JPO officials predict that such projects will lead to a large increase in the number of patent applications on human genes, including SNPs and cDNA clones. In fact, the Helix Research Institute, a genomics company, has already filed patent applications on more than 6,000 full-length human cDNA clones (see *Nature* 401, 3 & 520; 1999).

Helix, a joint venture between the Ministry of International Trade and Industry and ten private companies, plans to create a portfolio of intellectual property that can be shared by a planned consortium of 20 biotech companies and research institutes.

But many argue that Helix's patent applications, which largely consist of cDNA clones with unknown functions, are not valid, as they do not demonstrate the invention's utility.

JPO's guidelines on patenting human genes state that the gene must be sequenced and demonstrate utility, such as in diagnosis or treatment. They also specify conditions for patenting a sequence linked to a disease (that is, mutations such as SNPs), which can be used as a probe.

"We may not be successful with obtaining the patents, but the importance lies in the precedent that we have set for patent applications on full-length cDNA clones," says a spokesman for Helix.

Asako Saegusa

Research chair bonanza in Canada

Montreal

Canadian prime minister Jean Chrétien has announced that the federal government is to fund 2,000 new research chairs at universities, 1,200 of them over the next three years.

"This is incredibly exciting," commented Robert Prichard, president of the University of Toronto. Henry Friesen, president of the Medical Research Council, calls the announcement "stunning".

Chrétien, who has dismissed reports of large-scale emigration of scientists for better opportunities elsewhere, said in his announcement that the plan is "for brain gain, not brain drain".

The plan was proposed by the heads of the research fund-granting councils, the Canada Foundation for Innovation and some university presidents, in particular the rector of the Université de Montreal, Robert Lacroix, and the president of the University of British Columbia, Martha Piper.

The purpose, said Chrétien, is to "brand"



Chrétien: 'brain gain, not brain drain'.

Canada around the world "as the place to be for knowledge creation as we enter the twenty-first century", to enable Canadian universities to create research opportunities for the "best and the brightest Canadians", and to attract "some of the world's best minds" from other countries.

The prime minister failed to mention the cost of the scheme. But Friesen says the total cost of the 2,000 chairs over the first five years will be about Can\$300 million (US\$200 million).

Support for individual chairs, which could be renewed and revised after five years, will be Can\$100,000 for younger scientists and Can\$200,000 for more established researchers; most of this would be in salaries. ▶

- About 35 per cent of the total will go to the biomedical sciences.

"The scale of this is really quite remarkable," says Friesen. Taken together with previously promised increases, the biomedical research budget will be close to Can\$600 million in five years, Friesen says. "Not so long ago it was a Can\$230 million commitment."

The proportion of younger to senior research chairs has not been decided. A committee, including the three research granting council presidents and David Strangway, president and chief executive officer of the Canada Foundation for Innovation, will work out the plan's details.

Louis Siminovitch, university professor emeritus at the University of Toronto, sees the initiative as "a great thing", but says "it is not clear what the rules are going to be. Salaries are very important for recruiting people," he adds. "But that's only a very small part of the brain drain. What determines whether a person goes to the United States is research resources."

Siminovitch says the plan "is a very big step, but it may not be competitive with conditions in the United States. The most important message is that the federal government has awakened to the fact that they've got to do something for research and development."

Another investment of Can\$11 million has also been announced by Strangway. He says it will provide 124 researchers with the infrastructure necessary to conduct 64 research projects in 26 Canadian universities and colleges. Because the Canada Foundation for Innovation funds up to 40 per cent of a project, this investment may trigger a further Can\$16.5 million from other funding partners.

David Spurgeon

US university fears allayed over access to research data

Washington

Final regulations covering access to university researchers' data under the Freedom of Information Act were published in Washington last week. They seem to have gone some way towards satisfying both sides in a bitter dispute over how far such access rights should reach.

The regulations, published by the White House Office of Management and Budget (OMB), will allow raw data to be accessed under the act if these data have been used as the basis for a range of government actions.

But data behind unpublished research results, intellectual property and clinical records of individuals will be exempt. This aspect of the regulations addresses some of the research community's greatest fears.

"We feel that OMB has done a pretty good job — perhaps a better job than seemed possible," says Bill Colglazier, executive officer of the National Academy of Sciences, which has opposed the move to use the act to gain access to researchers' data.

George Leventhal at the Association of American Universities, which represents leading research universities, says: "We think OMB has done well, and we're especially pleased that commercial and personal data will be exempted."

The regulations have also been accepted by Senator Richard Shelby (Republican, Alabama), who championed the legislation ordering OMB to draw them up a year ago. Shelby had objected to the last draft of the



rules, published in August, saying that they allowed too many exemptions.

OMB acceded to Shelby's demand that the act should allow access to data behind all types of government action, not just federal regulations. It also removed a suggestion that the rules should apply only to research whose economic impact was expected to exceed \$100 million.

The final regulation "is a good first step to giving the American people access to the research and science used in federal policies that affect the lives of Americans each day", Shelby said in a statement.

US business has lobbied hard for access to research results under the act, seeing this as a means of assessing university studies used by the government as the basis for environmental and other regulations.

Colin Macilwain

French students continue to turn away from science

Paris

The number of students studying science in French universities and high schools has fallen for the fourth year in a row, according to figures released last week by the ministry for national education, research and technology.

Between 1995 and 1999, the number of first- and second-year university students — or students entering for a *diplôme d'études universitaires générales* (DEUG) — in the sciences fell from 150,000 to 126,400. The number of high-school students graduating with a science baccalaureate also dropped, from 143,200 to 125,000.

Officials say the reduced interest at universities stems from the rise in popularity of *classes préparatoires*, which feed into the prestigious *grandes écoles*, as well as shorter degree programmes at technical colleges.

Students in science DEUG programmes at French universities

Year	Number of students
1994-95	145,699
1995-96	149,688
1996-97	143,759
1997-98	134,447
1998-99	126,403

A survey distributed to 3,000 high-school students in April also shows that universities have a poor reputation among students, who think the quality of education is much lower than in the *grandes écoles*.

France is not alone when it comes to students straying from science. Statistics are similar elsewhere in Europe and the United States. The presidents of Europe's physics societies met last month in the United Kingdom to discuss the lack of students (see

Nature 401, 102; 1999). In Germany, for example, the number of first-year physics students has halved to 5,000 since 1991.

Physics appears to be the worst hit of the sciences. "Students are turning away from the 'hard' sciences, especially physics, and are more attracted by biology and astronomy," says Claude Feuerstein, president of the University Joseph Fourier at Grenoble. "It's not that it is difficult to find a job after studying physics, in fact it may even be easier than after completing a biology degree."

Feuerstein hopes that reforms, including an emphasis on experimentation over theory and even a marketing campaign, will boost the image of physics and the other hard sciences. "Students need an attractive context and to know why they are learning this."

Heather McCabe

PNAS joins peer-reviewed PubMed Central

Washington

The *Proceedings of the National Academy of Sciences* (PNAS) will appear on PubMed Central, the US National Institutes of Health's (NIH) new research archive, one month after the publication of the journal. But the academy's governing council will only allow this if research that has not been peer reviewed is kept off the repository.

Nick Cozzarelli, the editor of PNAS, says that the journal will be available on PubMed Central in January for a trial period of a year. He says that subscribers will have electronic access to papers two weeks before publication — six weeks before they become freely available to everyone else.

But the council of the National Academy of Sciences (NAS) said in a statement that “participation in PubMed Central is contingent upon its not including reports that have been screened but not formally peer reviewed”. The statement went on to say that non-peer-reviewed material, which the NIH has said could be submitted by any group of three grant-supported researchers, would have to be “completely separate” from peer-reviewed material submitted by journals.

“The council felt that making non-peer-

reviewed as well as peer-reviewed material available will confuse both scientists and the public,” says Ken Fulton, the executive director of the academy. He adds that the degree of separation “remained to be discussed,” and that “we haven’t had the chance to talk to NIH about what this means”.

But some officials at PNAS and the NIH are playing down the condition's significance, saying that PubMed Central will consist entirely of peer-reviewed material when it launches in January. David Lipman, director of the National Center for Biotechnology Information at the NIH and one of the main architects of PubMed Central, says that, although “we want to keep open the possibility” of publishing material prior to peer review, “all of our work is going into the peer-reviewed side” in the run up to the launch. “The non-peer-reviewed side is just not where the interest is.”

Lipman said that the NA statement was “complete consistent with our plans” and that any unreviewed mater

would appear under a different name, “such as PubMed Unreviewed”.

PNAS, a general-interest, biweekly journal with an international circulation of 10,000 and an emphasis on molecular biology, is the second major journal to join PubMed Central. *Molecular Biology of the Cell*, the journal of the American Society of Cell Biology, will also participate, as will half a dozen smaller journals.

Some geneticists, such as Pat Brown of Stanford University in California, advocate a preprint server for the life sciences community, analogous to the one that Paul Ginsparg runs for physicists at the Los Alamos National Laboratory in New Mexico.

Biologists are unaccustomed to the idea of wide circulation of preprints, and worry about the proliferation of junk medical science on such a server. The NAS council also said that it would form a committee “to consider the consequences of PubMed Central on science and science publishing”. **Colin Macilwain**



Funding changes aim to reform German universities

Munich

Germany's Bund-Länder Kommission (BLK), which negotiates joint financing of education and research between federal and state governments, will discuss next week the implementation of a new special fund for financing reforms in universities and other higher-education institutions.

The programme, to be launched in 2001, will increase the number of women in top academic jobs, promote the technically orientated higher education establishments known as Fachhochschule, and improve research infrastructure in east Germany. It will also earmark funds for improving information technology in universities and for modernizing university management and organization.

The latter moves are intended to encourage universities to become more competitive. They include paying professors according to performance and reorganizing faculties and departments to facilitate interdisciplinary teaching and research.

The programme will be the fourth in a series which, until now, have been called *Hochschulsonderprogramm* (HSP). Although *Länder* (state) governments are responsible for universities, the HSPs offer a mechanism for the federal government to

provide additional funding in areas it wants to promote.

In the past, federal and *Länder* governments have contributed equally to HSP financing. But the new Social Democrat-Green coalition federal government wants to drop the name HSP and introduce the measures on an annual basis, instead of the four- or five-year basis of previous programmes. It is prepared to provide DM420 million (US\$234 million) for 2001. It also wants to take over full funding of some of the measures designed to promote competition in universities.

A decision may not be reached next week, as *Länder* governments are reluctant to give the federal government full responsibility for any part of university policy, seeing it as an erosion of their constitutional rights.

A total of DM60 million per year is likely to be set for boosting the role of women in academic life, similar to the sum for this in HSP3, which expires next year. The money will be used to introduce gender studies more widely in universities, to increase the number of female students on scientific and technological courses and, most important, to increase the chances of women qualifying for top academic positions.

According to the plans put forward by the research ministry, less than 15 per cent of the new funds should be used to support PhD programmes, on the grounds that women are well represented at this level. The ministry proposes measures to increase the representation of women at higher levels, including the creation of short-term group leader positions to give women experience in leading research teams.

The ministry is proposing earmarking professorships for women by paying for the overlapping years of a number of duplicate professorships for women — an approach already used to provide posts for young researchers before the retirement of those occupying senior positions.

The federal research ministry did not want to extend the funding of research infrastructure in east German universities provided for in HSP3. But the eastern *Länder* have persuaded it that a significant gap still exists between east and west. The new programme now foresees DM20 million per year to help bridge the gap.

Josef Lange, general secretary of the German University Rectors' Conference, broadly welcomes the proposals but says that all HSP measures should lead to stable changes in universities. **Alison Abbott**

Genetic variations can point the way to disease genes

London

Researchers have shown for the first time that the search for human genes can be substantially narrowed using maps of small variations in the genome, according to a presentation this week at the American Society of Human Genetics (ASHG). They also claim that the density of variations needed is much lower than some have suggested.

The search for disease genes usually involves long sections of DNA, as traditional markers are widely spaced. An alternative is to use maps of single nucleotide polymorphisms (SNPs), based on the difference between individuals. Because SNPs are usually either part of a gene or close to a gene, they are seen as more effective than conventional methods for locating disease genes.

But some researchers have argued that a high density of SNPs would be needed to detect association between a marker and a disease. A study earlier this year based on simulations suggested that one SNP per 6,000 base pairs was needed — a total of 500,000 for whole-genome studies (*Nature Genetics* 22, 139–144; 1999). If true, says Allen Roses, director of genetics at Glaxo Wellcome, this would “put a damper” on the idea of researchers using SNPs because the technique is “tremendously expensive”.

But Roses claims that his work on other diseases has shown that an SNP map with a density of one per 10,000–30,000 base pairs could be used to identify areas of DNA thought to be linked to a disease — five times less dense than previously thought.

Roses told the ASHG at its meeting in San Francisco this week that Glaxo Wellcome scientists have already used SNPs to locate areas of DNA for three diseases: migraine; type II diabetes; and psoriasis. Roses also says he has shown that, given a known susceptibility gene, APOE in Alzheimer's disease, SNPs can narrow its location down from a strip of DNA four million base pairs in length to only 40,000 base pairs — a region containing only two genes.

“It is important to demonstrate that it is possible to narrow the region by orders of magnitude,” says Roses. “Nobody has done SNP mapping before because it is tremendously expensive and unproven. We felt it important to demonstrate to the academic world that it is possible.”

Natasha Loder

NASA plans to map stars and hunt gamma-ray bursts ...

Washington

The US space agency NASA last week gave the go-ahead to two new medium-class Explorer spacecraft in the next five years. One will study gamma rays, and the second will map 40 million stars with unprecedented precision.

The \$163-million Swift Gamma-ray Burst Explorer, scheduled to reach orbit in 2003, will have three telescopes — viewing in gamma-ray, X-ray and ultraviolet/visible wavelengths — poised to observe these mysterious bursts within minutes of their appearance.

During its three-year mission, Swift is expected to pinpoint some 300 gamma-ray bursts, and to discover about 400 black holes. The principal investigator is Neil Gehrels of NASA's Goddard Space Flight Center in Maryland.

The \$162-million Full-sky Astrometric Mapping Explorer (FAME), due to launch in 2004, is intended to improve on the successful European Hipparcos mission of the early 1990s, which gave extremely accurate positions (to within 1 milliarcsecond) for 120,000 stars.

FAME will use advanced charge-coupled device arrays and a more efficient scheme for blocking out sunlight to pinpoint some 40 million stars down to fifteenth magnitude, with an accuracy of 500 microarcseconds. For stars brighter than ninth magnitude the accuracy will be 50 microarcseconds.

“There's all kinds of science you can do with these observations,” says Kenneth Seidemann, leader of the project's science working team at the US Naval Observatory in Washington DC.

A key objective is the search for wobbles in a star's motion that betray the presence of a large planet or brown dwarf. Seidemann says FAME should be able to detect planets that are as small as twice the mass of Jupiter.

During its five-year, all-sky survey FAME will produce a catalogue of star locations that can be used by NASA's planned follow-on, the Space Interferometry Mission, which will be accurate to one microarcsecond. The European Space Agency is also considering an astrometry mission called GAIA, which could catalogue a billion objects with similar precision.

Tony Reichhardt

... and Europe to measure the Earth's gravity

Munich

A major mission to determine the Earth's gravity field at high resolution has been recommended by Earth scientists as top priority in the European Space Agency's (ESA's) Earth Observation programme, 'Living Planet'.

The mission, called GOCE (Gravity Field and Steady State Ocean Circulation Explorer), was selected by ESA's Earth Sciences Advisory Committee in a competition held in Granada last week between four proposals developed by the European scientific community.

An Atmospheric Dynamics Mission (ADM) — a laser system for measuring wind and wind profiles from space — was the runner-up in the competition.

Both missions are expected to be approved by ESA's decision-making



Up and away: Antarctica's Larsen Ice Shelf from space.

Science Programme Committee next month.

They should also provide important data for climate research.

GOCE will give information on irregularities in the mass

distribution of the Earth by measuring variations in the gravity field.

This, together with satellite altimetry of the oceans' surface, will enable oceanographers to determine ocean circulation in more detail. They will also be able to assess the variation in the polar ice caps more accurately.

By measuring winds at different heights, ADM will help to clarify how the atmosphere moves. GOCE will also improve estimates of the density of the Earth's lithosphere and upper mantle.

GOCE is set for a launch in mid-2004, and ADM in 2006. Both missions will cost less than 400 million euros (US\$435 million) and will be run by scientific teams consisting of members from most of the ESA's member states.

Alison Abbott

ESA

US Senate ignores scientific advice in failing to ratify test ban treaty

Researchers presented data that supports signing the test ban treaty, but they found that most politicians had already made up their minds.

Washington

When the US Senate announced on 1 October that it would vote on the Comprehensive Test Ban Treaty within 12 days, senior researchers with a stake in the treaty wasted little time.

Frank Von Hippel, professor of public and international affairs at Princeton University, and Ray Kidder, a physicist at the Lawrence Livermore National Laboratory in California, travelled to Washington early the following week, seeking moderate Republican 'swing senators' whom, they thought, would determine the outcome of the debate. They were in for a surprise. "We couldn't find any," Hippel laments. "The result was already stitched up."

Others lobbying for treaty ratification were similarly frustrated. Two key technical issues underpin the US debate: whether seismologists can verify that other countries do not cheat, and whether the scientific stockpile-stewardship programme can ensure the safety and reliability of US nuclear weapons in the absence of testing.

To make the first point, the American Geophysical Union and the Seismological Society of America issued a joint statement on 6 October declaring that the treaty's proposed monitoring system "can be relied upon" to detect cheating.

Interpretation of data

But by the time the statement was issued, the argument was already lost. On Sunday 3 October, the lead story in the *Washington Post* announced that "the Central Intelligence Agency [CIA] has concluded that it cannot monitor low-level nuclear tests by Russia precisely enough to ensure compliance" with the treaty.

The story contained no attributable quotes, and no response from US seismologists, who have been arguing with CIA analysts for years about the interpretation of seismic data from the Novaya Zemlya test site in northern Russia.

The CIA repeated this assessment in closed briefings of senators on Tuesday 5 October. Although friends of the treaty say that no fresh information was revealed, its enemies cited their classified content — which, of course, they could not divulge — as proof that the treaty was unverifiable.

As for the effectiveness of stockpile



Testing time: the treaty would ban nuclear tests like this 1995 French one at Fangatau Atoll.

stewardship, much technical advice also favoured ratification. Thirty-two US Nobel prizewinners in physics signed a letter to the Senate, stating that "fully informed technical studies have concluded that continued nuclear testing is not required to retain confidence in the safety, reliability and performance" of US nuclear weapons.

But none of the Nobellists testified before the Senate. The experts who did were the directors of the three nuclear weapons laboratories: Los Alamos and Sandia in New Mexico and Lawrence Livermore. Their testimony before the Armed Services Committee on 7 October was finely balanced on the question of stockpile stewardship, although John Warner (Republican, Virginia), who chairs the committee, believes it cast doubt on the programme's effectiveness.

"On balance, we remain positive about the prospects for success," said Bruce Tarter, director of Livermore, whose testimony was perhaps the most favourably disposed towards stockpile stewardship. "With sustained support, it is an excellent bet but it ain't a sure thing."

Paul Robinson, the director of Sandia, came closest to repudiating the programme's ability to maintain the weapons stockpile indefinitely. "The difficulty we face is that we cannot guarantee that science-based stockpile stewardship will be ultimately successful; nor can we guarantee that it will be possible to prove that it is successful."

The three directors had a tough day on 7 October. It was time to choose between pleasing their boss, energy secretary Bill Richardson, or their paymaster, the Republican majority in Congress.

But treaty supporters believe that the director's testimony hurt its already fading prospects. "The lab directors' testimony was

misdirected," says Spurgeon Keeney, president of the Arms Control Association. "They behaved badly, and used the hearing as an opportunity to plead for their budgets."

Daryl Kimball, head of the coalition lobbying for the treaty in Washington, said: "They tried to engage in a game of budgetary extortion, by insisting that they weren't 100 per cent confident in stockpile stewardship."

It was left to Sid Drell of the Stanford Linear Accelerator Center in California, one of the architects of the stewardship programme, to defend it before the Armed Services Committee. "The coin of the realm in science is not opinion, but data," he said. "Any scientist welcomes more data. The question is what data are necessary." Nuclear explosions, he testified, would contribute "nothing essential" to the maintenance of the nuclear stockpile.

However, like most of the technical witnesses, Drell was speaking to an empty room at the end of a long day. Only Warner and Carl Levin (Democrat, Michigan) were there to hear him — and both were having whispered conversations with aides.

Party politics

The whereabouts and motivations of the phantom Republican moderates are difficult to pin down. Warner, whose early declaration against the treaty virtually ended its prospects of ratification by the necessary two-thirds majority, seemed preoccupied with his concern that "unsafe" nuclear weapons might blow up in the hands of "young servicemen and women".

Richard Lugar (Republican, Indiana), the moderate Republican whose support the treaty most badly needed, kept a low profile, refusing to meet with scientists or talk to the press and issuing a comprehensive rebuttal of the treaty on 8 October.

Pete Domenici (Republican, New Mexico) wanted the vote deferred, and came close to saying the treaty should be ratified with some adjustments, before voting against it.

In the end, partisan politics seems to have triumphed over technical debate. After the treaty fell, by 51 votes to 48, Clinton emerged energized, doing what he does best — attacking the Republicans. Vice-president Al Gore was also on television, with the first commercial of his 2000 presidential campaign, in support of the treaty.

Colin Macilwain

US guidelines widen the net on scientific misconduct

London President Bill Clinton's administration last week proposed new guidelines for investigating scientific misconduct involving government-funded research.

The long-awaited guidelines, issued by the White House's Office of Science and Technology Policy, propose extending policies beyond the federal agencies with the most misconduct experience — the National Institutes of Health and the National Science Foundation (NSF) — to other agencies that fund research.

And, while keeping the traditional definition of misconduct, which includes fabrication and falsification, the guidelines propose that plagiarism in published research should also be considered as misconduct.

The administration's proposals were developed through the National Science and Technology Council, a cabinet-level group that develops presidential policies for federal agencies. The procedures have been published in the *Federal Register* for a 60-day comment period, before the implementation of final guidelines.

The National Academy of Sciences is to

hold a meeting on the guidelines on 17 November in Washington, sponsored by the NSF and other agencies.

Russian institute challenged over seized fossils

Munich A lorry-load of fossils seized by customs officials at the Russian–Finnish border last December has been found by Russian palaeontological experts to contain many scientifically valuable samples, including two previously undescribed species of bear.

Many of the items had originally been described by the directors of the Russian Academy of Sciences' Institute of Palaeontology (PIN) as having no scientific value, so the ministry of culture had issued export licences. The items will now be offered to Russian museums, and government officials have stripped the PIN of its formal advisory role to the ministry.

Experts fly to Japan's nuclear accident site

Tokyo Experts from the International Atomic Energy Agency (IAEA) arrived in Japan last week to inspect the nuclear accident site in Tokaimura. Their arrival coincides with an announcement by the Science and Technol-

ogy Agency (STA) that it may upgrade the severity of the accident, which exposed 69 people to high levels of radiation, from level 4 to level 5 on the international scale.

Although the Japanese government declined IAEA's initial offer to send experts to the uranium-processing facility when the accident took place on 30 September, the STA decided to accept their offer "to regain confidence [in Japan's nuclear safety] among the international community". A team of nuclear experts from the US Department of Energy also arrived in Japan on Monday (18 October) to inspect the accident site.

Mobil buys into Russia's oil production know-how

Moscow The Mobil Technology Company has signed a US\$330,000 agreement with the All-Russian Institute for Theoretical Physics (VNIITF) in Snezhinsk — the nuclear city formerly known as Chelyabinsk — and two institutes of the Russian Academy of Sciences. The agreement addresses the modelling of oil flow in porous media, and will provide Mobil with sophisticated mathematical techniques for optimizing oil production.

The agreement brings one of the earliest investments in VNIITF by a commercial organization, and the institute's first investment through the Partner Programme

of the Moscow-based International Science and Technology Centre. It brings together VNIITF scientists with those of the Institute of Mathematical Modelling and the Institute of Numerical Mathematics.

Fears for NIH funding as Porter plans to go

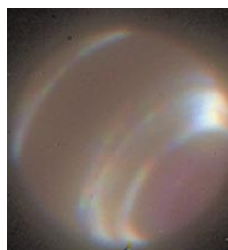
Washington John Porter (Republican, Illinois), chairman of the appropriations subcommittee that funds the US National Institutes of Health (NIH) and biomedical research's most powerful advocate in the House of Representatives, is to retire from Congress at the end of next year.

Porter's surprise announcement was believed to be partly motivated by Republican rules that would require him to move on from chairmanship of the labour, health and education subcommittee at that time. Porter's departure, following the resignation this month of Harold Varmus as NIH director, could make it tougher for the agency to sustain the sharp funding increases it has enjoyed in recent years.

Cameras on Earth match Hubble's Neptune images

London Some of the sharpest ground-based images of the planet Neptune have been cap-

tured with cameras using adaptive optics. New cameras have been used at two observatories, with the best images of Neptune taken from the ground so far being achieved at the Keck Observatory in Hawaii (above). Adaptive optics adjusts a secondary mirror to compensate for distortions caused by the Earth's atmosphere and to allow ground telescopes to produce images as sharp as those from the Hubble Space Telescope.



Germ warfare inquiry to quiz servicemen

London An inquiry into the death of a man following UK germ warfare experiments at Porton Down chemical warfare centre in 1953 is to be widened to include injuries and deaths to a number of other servicemen. The original investigation, announced in August, was confined to the death of a Royal Air Force aircraftman, Ronald Maddison (see *Nature* 400, 809; 1999).

The investigation was launched after Wiltshire police received complaints from a former serviceman who participated in the tests. The investigation has grown as more

relatives have come forward to the police with complaints.

The experiments were conducted at Porton Down to determine how much of the nerve gas Sarin was needed to penetrate a military uniform. According to press reports, 300 former servicemen who claim their health has suffered because of taking part in the experiments are expected to be interviewed by police.

Kansas education chair refuses Ig Nobel Prize

Washington At a meeting of the Kansas State Board of Education last week, board chairwoman Linda Holloway declined the Ig Nobel Prize offered by University of Kansas geneticist Douglas Ruden. The Kansas and Colorado education boards were granted the Ig prize at Harvard University last month for "mandating that children should not believe in Darwin's theory of evolution" (see *Nature* 401, 518; 1999).

Ruden claims that Kansas' new creationist standards have backfired, only serving to spur interest in evolution. At the same meeting, Holloway did accept a 'consciousness-expanding' cookie offered by local Hare Krishna leader Danavir Swami, who staunchly maintains that man did not evolve from apes.

Millions at risk as big cities grow apace in earthquake zones

Sir — The recent earthquakes in Taiwan, Turkey and Greece are a reminder that the world's urban population is increasingly vulnerable to earthquakes. About one-third of the world's supercities — those with populations of more than two million — are located near tectonic plate boundaries, where damaging earthquakes have occurred and will recur.

Between 1950 and 2050 the world's population will have increased from 2.5 billion to 8.9 billion, most of whom will live in great cities. In 1950 there were two megacities with populations of more than eight million, whereas there are now 27. All cities close to plate boundaries, and many that are not, are vulnerable to earthquake damage some time during the next millennium. Although earthquakes with more than a million fatalities have not occurred in human history, the availability of many new urban targets with populations approaching tens of millions makes such an event possible during the next century.

Offsetting this bleak view of potential fatalities is the recognition that the additional three billion people anticipated in the next 50 years will require a 50 per cent increase in urban construction. Although half of this future growth is expected in parts of Africa where earthquakes are fortunately rare, much future growth is expected in Asia where rates of seismic activity are high. Loss of life from future earthquakes could be much reduced if new construction in known seismogenic zones was designed to withstand shaking.

The reconstruction costs following an urban earthquake have risen to levels that represent an economic burden not only on local economies, but also on the global economy, making the incorporation of earthquake resistance economically attractive.

The absence of earthquake resistant construction in future cities would be indefensible. But the issue may remain a low priority in many developing nations.

Roger Bilham

Department of Earth Sciences, University of Oxford, Parks Road, Oxford OX1 3PR, UK

Where are the high-tech entrepreneurs?

Sir — Craig Pickering argues that a culture of venture capitalism must be created in Europe (*Nature* **401**, 209–210; 1999). I do not believe that lack of government

support or venture capital is the reason for the undynamic high-technology landscape, in Germany at least. The problem is a lack of entrepreneurial spirit.

As someone who has recently founded a company with initial finance of almost DM15 million (US\$8.2 million), my impression is that there is hardly a country where start-up venture capital is so easily accessible as it is here — it is much easier to obtain in Germany than in the United States.

At capital fairs, such as the recent Innovation Days in Berlin, there are dozens of banks and venture capital companies exhibiting but hardly any would-be entrepreneurs. The impression left is one between ridiculous and pathetic.

Germany frequently has publicly organized business plan competitions. And the government understands the overwhelming importance of high technology for the future of the economy. Politicians here are showing enthusiasm and respect for high-technology entrepreneurship.

The European problem cannot be solved by just money or political support. It is a problem of culture. I began to develop the vision for my company eight years ago. A feasible scientific concept evolved during my undergraduate and PhD studies. What was needed was a constant belief that at some point a business would be made out of all this. Once the money seemed available, I was well prepared to jump at it. You need to wish to be an entrepreneur over years of your scientific career. But few Europeans grow up with the desire to found their own businesses. Now that everybody tells them to do so, they neither have the heart nor mature enough ideas.

We Europeans must hope that being a venture capitalist in Europe is not so frustrating that private equity sources will have dried up once today's — hopefully more entrepreneurial — generation of students tries to commercialize its ideas.

Alexander Olek

Epigenomics, Kastanienallee 24, 10435 Berlin, Germany

The birth of Big Biology

Sir — Your editorial "Big Science comes of age"¹ suggests that, in my writings of 40 years ago, I neglected Big Biology. On the contrary, I argued that "We are entering an age of biomedical science and biomedical technology that could rival in magnitude and richness the present age of physical science and physical technology"².

In arriving at this view I was much influenced by the biologist-cum-engineer Norman G. Anderson, the inventor of the zonal centrifuge. Anderson proposed what

he called the Molecular Anatomy programme³ — cataloguing and characterizing the 100,000 or so human proteins. Part of Anderson's programme was undertaken at Oak Ridge National Laboratory, but a serious, systematic effort has had to await the success of the Human Genome Project (which, incidentally, he had also proposed). Both projects represent Big Biology; both owe much to Anderson's prescience.

Alvin M. Weinberg

Oak Ridge Associated Universities, PO Box 117, Oak Ridge, Tennessee 37831-0117, USA

1. *Nature* **400**, 387 (1999).

2. Weinberg, A. M. *Minerva* **IV**, 5–14 (1965).

3. Anderson, N. G. *Science Journal* January, 1–7 (1967).

Corot would put more planets in the picture

Sir — Declan Butler reports on the difficulties the French satellite Corot is facing owing to budgetary reductions (*Nature* **401**, 202; 1999). This report is mainly correct, but I must respond to comments made by David Hughes and cited in the report.

According to the report, Hughes argues that "extrasolar eclipses are rare, even when one is looking at 6,000 stars" and that "by definition, an eclipse would also be irreproducible".

These comments are inappropriate for the following reasons. First, Corot will observe at least 30,000 stars, not 6,000, as there will be five target fields, each observed for 150 days, and each with a required minimum of 6,000 stars. The exact number depends on the density of stars in the chosen target fields and could increase by up to $5 \times 10,000$.

Second, ground-based detection of extrasolar planets has shown that at least three to five per cent of stars have giant planet companions. The geometrical probability of having an eclipse by a planet closer than 0.3 AU is 1.5 per cent. So, over five samples of 6,000 stars, and assuming a similar planet/star ratio to that applied to terrestrial planets, Corot should detect a total of 13 to 22 such planets.

Third, a planet with an orbital period of less than 50 days would, in general, make at least three eclipses in each 150-day run, making them, "by definition", reproducible.

Finally, single-eclipse events will also be analysed, thanks to the two-colour system implemented on the exoplanet channel, which will help reduce false alarms. About ten such events could be detected in this way.

Annie Baglin

Corot Team, DESPA, Observatoire de Paris, 5 place J. Janssen, 92195 Meudon, France

What is a wave?

Most people assume that a wave, being central to all the phenomena we observe, has a uniform definition. But defining this basic concept isn't so easy.

John A. Scales and Roel Snieder

Ask your colleagues and students to define a wave, and you may be surprised at the answers you get. Even wave professionals are prone to confusion and vagueness when confronted with such an apparently simple question. Students often begin circularly: "a wave is a solution to the wave equation." But what is a wave equation? Professionals are more likely to mutter something about propagation speed, as if smell and heat didn't propagate with some speed. Mathematicians tend to give formal characterizations based on the hyperbolicity of certain differential equations.

Just as a definition of noise must be grounded empirically, so to define a wave we should look at what nature has to offer. A preliminary answer might be: a wave is a propagating imbalance. The imbalance concept is also present in a simple oscillator where kinetic energy and potential energy are interchanged during the oscillation. Hamilton's principle (that the path taken by a dynamical system is the one that minimizes the time integral of the difference between the kinetic and potential energies) is a formalization of this idea of interchange between the two forms of energy. However, what makes a wave different from a single oscillator is that this imbalance propagates. (Of course, two travelling waves can form a standing wave, but let us ignore this complication for the moment.)

Stable equilibria

At this simplest of levels, the ubiquity of (classical) waves can be attributed to nature's love of stable equilibria. Whatever the forces that connect bits of matter together (electromagnetic or gravitational, for instance), for small perturbations about a stable equilibrium point, the forces are approximately linear; a linear restoring force implies harmonic oscillation; and coupled systems of oscillators support both propagating and standing disturbances. Linearity also implies superposition, so we can carefully add periodic solutions together to get finite wave 'packets'. So, for small perturbations about an equilibrium state of coupled or extended systems, waves are the natural consequence of the stability of simple harmonic motion.

The miracle of the waves that we see is the organization they display, but there are examples where this organization is destroyed. Strong scattering leads to diffusive

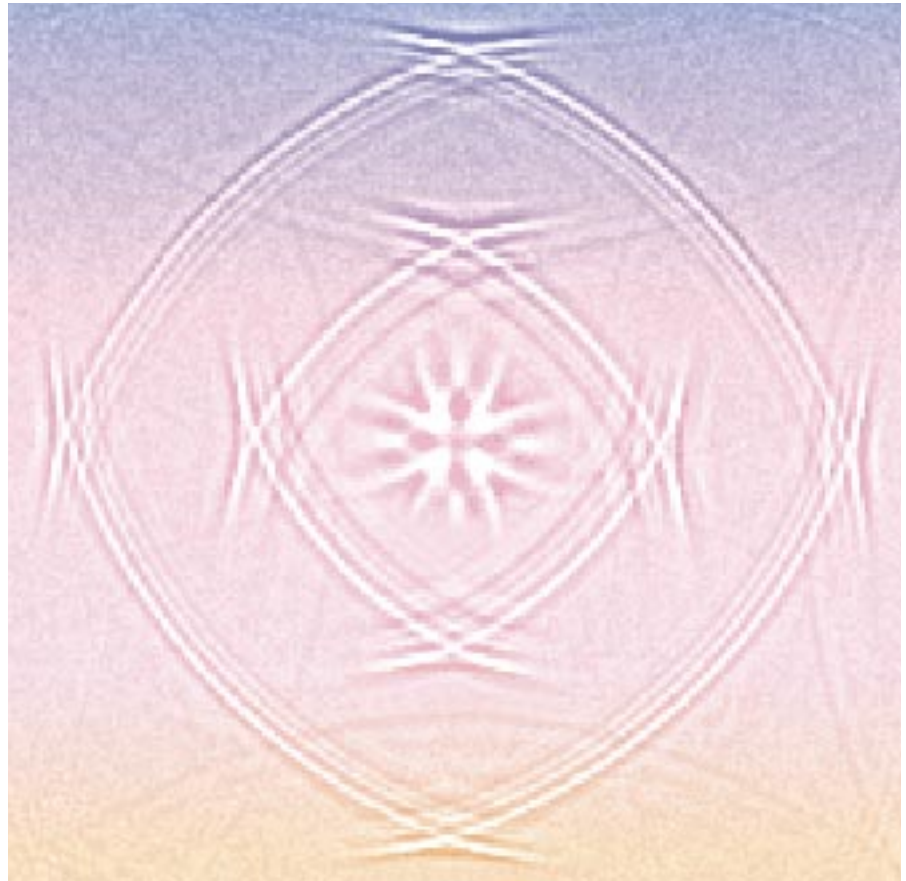


Figure 1 The recorded wavefield that has propagated through a homogeneous silicon crystal at three times after excitation at the other side of the crystal. Details of the experiment are given by Wolfe³. This example shows that wavefronts in a homogeneous medium can be square rather than round. (Courtesy of J. P. Wolfe.)

behaviour rather than wave propagation. The scatterers destroy the level of organization in the incident wave and ultimately lead to diffusive (un-wave-like) behaviour. Similarly, when a wave breaks on a beach, the advective terms in the equation of motion couple all the different length scales in the wave, and the organization we see in the swell is destroyed. Ultimately, the wave is dissipated as heat. In view of these examples, where a wave ceases to be a wave because of the destruction of its degree of organization, we are led to modify our definition of a wave to become: a wave is an organized propagating imbalance.

Wave propagation is in many situations described by a linear differential equation. In reality, nonlinearity is important, and this nonlinearity may destroy the waves. As an example, consider the waves on the beach again. Look far out at the ocean from any beach, and you will see ripples on the surface

of the water with a period of 5–10 seconds. As these ripples approach the beach, their heights increase until they can no longer support their own weight and they break catastrophically. Mathematically, this is caused by the nonlinear terms in the equation of motion becoming increasingly important as the waves grow. When nonlinearity becomes important, organized wave motion changes into turbulent motion. In this process, it is impossible to state exactly at which point the wave ceases to be a wave.

Nonlinearity is sometimes essential for maintaining the organization of a wave. In solitons, the wave spreading by dispersion is exactly (and miraculously) offset by the nonlinear steepening of the wave, so that a solitary wave maintains its identity. This means that nonlinearity can lead to the creation of organization as well as to its destruction.

The simplest soliton to produce is the

cylindrical hydraulic jump. Go into the kitchen and turn on the tap at the kitchen sink. As the water strikes the sink, its vertical momentum is converted into horizontal momentum. In most cases, when the water first hits the surface, it is travelling faster than the speed of surface-water waves, so disturbances cannot propagate as surface waves and are swept downstream by the water. But the water must slow down, and at some point it slows to the speed of surface-water waves. What happens then is truly remarkable. A jump, or shock, develops — the thickness of the water increases almost discontinuously. Further downstream, the water's surface is awash with surface waves that are now free to propagate. But the jump is stationary, so why should we regard this as a wave-like phenomenon? Well, imagine you are in a boat being swept downstream by the water. In your frame of reference the hydraulic jump is a solitary 'wave' racing upstream, much like a tidal bore.

Diffusion

To return to the original question, many people may say something about wave-like versus diffusive behaviour. In many physical phenomena, there is no clear distinction between these two extremes of behaviour. For instance, take light propagating in a turbid medium such as milk. The turbidity is the result of scattering. (The absorption cross-section of the fat molecules in the milk is much smaller than the scattering cross-section — the opposite of ink, for instance.) The equations governing the electric field are still the same linear-wave equations that follow from Maxwell's equations. But what the eye registers is not the electric field itself, but rather the intensity of the field. Because the field reaching the eye is the superposition of the uncountably many scattered waves originating in the milk (the equations are linear), the actual intensity is the intensity of this superposition. It is easy to see that this total intensity has both a coherent and an incoherent term in the superposition.

If the different scattered waves do not interfere constructively with one another, then the total intensity is merely the sum of the intensities of the individual waves. If these individual, non-interfering scattering terms are thought of as representing a vast number of uncorrelated brownian paths through the milk, it is no surprise that the no-interference intensity satisfies a diffusion equation (which is the equation of the probability distribution for brownian motion). So the electric field satisfies a wave equation, as Maxwell said it must, but the quantity we measure (the intensity) satisfies a diffusion equation.

Some phenomena are clearly diffusive, with no wave-like implications — heat, for instance. We all 'know' that heat conduction is governed by the diffusion equation. Maxwell actually had his doubts about this (see ref. 1 for a fascinating account). The



Even the ripples of space-time are waves.

standard diffusion equation doesn't take into account any propagation speed, so it cannot really be a fundamental description of the transport of heat; according to this equation, if you apply a heat source to one end of a rod, the temperature at the other end begins to change instantaneously! Maxwell, working from kinetic theory, imported a ballistic term into the equations of heat conduction. He ended up with the telegraph equation (it has first and second derivatives with respect to time), with its trade-off between the diffusive behaviour (which comes from the first time derivative) and ballistic behaviour (coming from the second time derivative). Maxwell dropped this ballistic term after concluding² that it "may be neglected, as the rate of conduction will rapidly establish itself".

That was consistent with experiments 100 years ago, but not any longer. As far back as the 1960s, ballistic heat pulses were observed at low temperatures. The idea is that heat is just the manifestation of microscopic motion. Computing the classical resonant frequencies of atoms or molecules in a lattice gives numbers of the order of 10^{13} Hz, that is, in the infrared, so when molecules jiggle they give off heat. These lattice vibrations are called phonons. Phonons have both wave-like and particle-like aspects. Lattice vibrations are responsible for the transport of heat, and we know that heat is a diffusive phenomenon. However, if the lattice is cooled to near absolute zero, the mean-free scattering path of the phonons becomes comparable to the macroscopic size of the sample. When this happens, lattice vibrations no longer behave diffusively but are actually wave-like. By controlling the temperature of a sample, one can control the extent to which heat is ballistic (wave-like) or diffusive. In essence, if a heat pulse is launched into such a sample (by passing a current through a wire, for instance), and if the phonons can get across the sample without scattering, they will propagate like waves. The more they scatter, the more diffusively they behave. When it's very cold, heat waves propagate as waves. Figure 1 shows an example; many more are found in ref. 3.

Waves are not only elusive in their character, their presence is ubiquitous in nature. Our two main senses, vision and hearing, rely

on waves. We call these the 'main' senses because they give us the most precise information about the environment. It is typical that there are common-language words for the loss of eyesight or hearing, but not for the loss of sense of smell, taste and warmth. Most of what we know about the world around us we learn through waves. In addition, neurons work by the propagation of electric waves through the axons. A prime example is the triggering of the heart by a propagating electric pulse through the heart tissue.

Even the ripples of space-time are waves. These are called gravitational waves and propagate at the speed of light. The first observatories with a real chance of detecting gravitational waves are now coming online. These instruments, enormously long interferometers, will be able to measure strains on the order of 10^{-20} or smaller.

Quantum mechanics

Another field where waves have a central role is quantum mechanics, from which we learn that everything has a wave character. Einstein used the relation $E = hf$ (energy equals Planck's constant times frequency) to connect the wave frequency of light with the energy of light's discrete quanta (photons). De Broglie extended this to electrons and other ponderable matter. For classical waves, dissipation generally damps the wave motion, and ultimately everything seems to come to rest. Quantum mechanics shows that matter waves do not suffer from dissipation. Even the ground state of the harmonic oscillator is in harmonic motion. Matter waves never come to rest. Taking this last idea a step further, one can conjecture that the ubiquity of waves is crucial to our concept of time. Change is the manifestation of time, and regular oscillations are a clear manifestation of change. Appropriately, the waves that propagate in quartz crystals are now the dominant tool used to keep track of time.

It is clear that waves are ubiquitous in nature and that they are central to the structure of matter and time as well as to many physical, biological and chemical phenomena. It is striking that the concept of waves is so hard to define, and that the distinction between wave-like and non-wave-like behaviour can be so fuzzy. Taking all these examples into account, we stick with our definition of a wave as an organized propagating imbalance; just don't ask us to define 'organized'.

John A. Scales is in the Department of Geophysics and the Center for Wave Phenomena, Colorado School of Mines, Golden, Colorado 80401, USA.
e-mail: jscales@mines.edu

Roel Snieder is in the Department of Geophysics, Utrecht University, PO Box 80.021, 3508 TA Utrecht, The Netherlands.
e-mail: snieder@geo.uu.nl

1. Maddox, J. *Nature* **338**, 373 (1989).
2. Joseph, D. D. & Preziosi, L. *Rev. Mod. Phys.* **61**, 41–73 (1989).
3. Wolfe, J. P. *Imaging Phonons: Acoustic Wave Propagation in Solids* (Cambridge Univ. Press, 1998).

Giants' footprints in the greenhouse

The seeds of our understanding of global warming were sown by early heroes.

Greenhouse: The 200-Year Story of Global Warming

by Gale E. Christianson
Constable/Walker: 1999. 305/316 pp.
£9.99 (pbk)/\$25 (hbk)

Historical Perspectives on Climate Change

by James Rodger Fleming
Oxford University Press: 1998. 208 pp. \$45,
£32.50

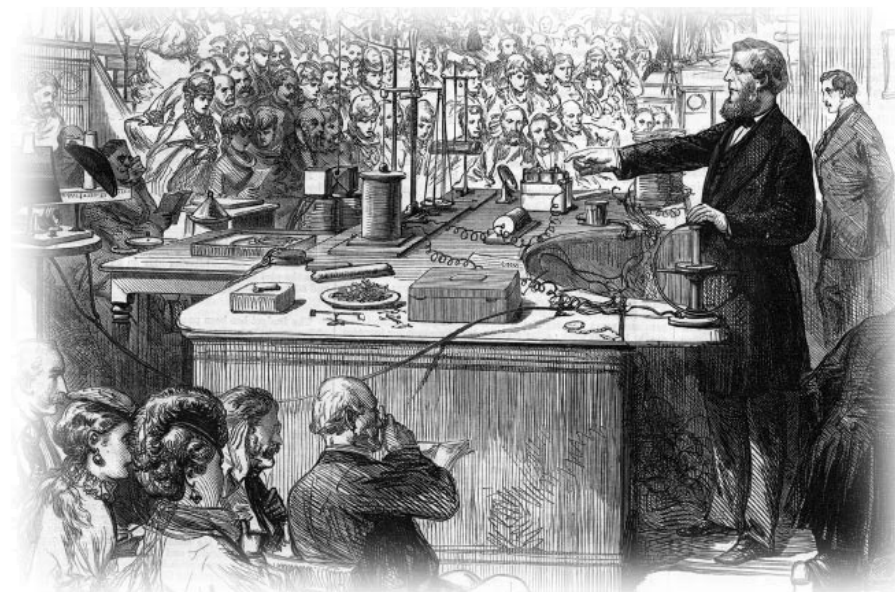
Robert J. Charlson

It is a rare opportunity for a scientist to be offered a chance to review the work of historians; more often, it is the other way around. In this case, two professors of history have written about the development of our scientific understanding of global warming. My interest is in their choices of topics and their descriptions of theories of the Earth's climate.

There is a growing awareness that human activities have changed the chemical composition of our planet's atmosphere and that climatic consequences seem imminent. The authors of both books, writing for non-meteorologists, say that the theory of the 'greenhouse effect' is not new, but has old and well-developed roots. Most of this theory was founded by scientific giants, notably Joseph Fourier, John Tyndall, Samuel Langley and Svante Arrhenius. Between them, these men recognized that the Earth's surface is some 30 °C warmer than it would have been without atmospheric absorption of radiation, and why it is so.

In the early nineteenth century, Fourier proposed that the atmosphere might "augment" Earth's temperature because ... "light finds less resistance in penetrating the air, than in repassing into the air when converted into non-luminous heat". And his ideas on heat diffusion are a precursor of today's notion of steady-state heat balance: heat "dissipated ... at the surface of the Earth is compensated at each instant by that received by the Sun".

Then, in the 1860s, Tyndall showed that gases can absorb and emit radiant heat. This, along with Langley's data on heat transmission through the atmosphere (made in the 1880s by measuring heat fluxes from the Moon), became the basis for the first heat-balance calculations by Arrhenius, published in 1896. His real quest was to explain the causes of the ice ages, but his agonizingly lengthy numerical calculations yielded the first estimates of the sensitivity of the Earth's surface temperature to changes in carbon dioxide. These estimates are amazingly close to those obtained by today's most advanced



Radiating knowledge: a lecture by John Tyndall, whose work on gases contributed to the climate jigsaw.

models. Arrhenius' work still stands as a paradigm for studies of the Earth system, and his use of the word 'hothouse' is an ancestor of our 'greenhouse effect' terminology.

Fleming carries the story through the debate, largely by geologists, about how the carbon dioxide hothouse hypothesis explained the differences between temperatures in interglacial and ice-age periods.

During the 1930s and 1950s, Guy Callendar first identified and attempted to quantify the anthropogenic increase in atmospheric carbon dioxide. He also argued that the 0.4 °C temperature increase in the decades just before 1950 was due to carbon dioxide. His estimate of the preindustrial level of 280 parts per million, while terribly uncertain, almost exactly matches today's best data from studies of ice cores.

Ever more accurate calculations of the transfer of infrared radiation from the Earth's surface and through the atmosphere by Gilbert Plass from the mid-1950s showed that increased carbon dioxide meant increased heating of the lower atmosphere. Then Roger Revelle showed theoretically why added carbon dioxide did not simply all dissolve in slightly basic sea water. At about the same time, Hans Suess showed, from studies of the $^{14}\text{C}/^{12}\text{C}$ ratio in tree rings, that carbon dioxide from fossil fuels had indeed built up in the first half of the twentieth century. Charles Keeling's attempts at continuous observations of carbon dioxide in Antarctica, showing a clear increase in carbon dioxide, were first published in 1960.

Keeling's data recorded at Mauna Loa Observatory, Hawaii, have become the hall-

mark of studies of anthropogenic enhancement of the greenhouse effect, showing conclusively that carbon dioxide has indeed increased by around 30% since the middle of the nineteenth century.

Neither author discusses the many developments of the past 20 years. For example, although both mention the controversy in the mid-1970s that global cooling might be caused by a build-up of anthropogenic aerosols (dust, smoke and haze), neither continues to the current understanding that such man-made particles do cause significant loss of solar irradiance. They also omit the landmark papers by Claude Lorius beginning in 1985, which clearly show low carbon dioxide levels strongly correlated with low temperatures in the last ice age. (Arrhenius would have been delighted.)

The two books are very different in tone and presentation. Fleming presents the more academic work, with a superb bibliography, and places the development of the science of the greenhouse effect in a context of the scientific concepts of the day. Christianson, on the other hand, places developments in the broader human context of the industrial revolution and the rise of environmentalism. In doing this, he comes precariously close to advocacy of an activist agenda. He also seems to accept as proven that the temperature increase of the past century was caused by increased greenhouse gases, and that the science is a 'done deal', without objectively giving an accurate picture of uncertainties.

One subtle but important aspect missed by both is the shift in assessing global warming from calculating the temperature change

to estimating the forcing (imposed change in heat balance, expressed in watts per square metre) caused by the change in atmospheric composition. Before the mid-1980s virtually all estimates of the greenhouse effect were given as temperature changes. Recognition that there are many greenhouse gases, as well as aerosols, that affect the Earth's heat balance required ways of assessing and comparing their effects.

This led Robert Dickinson and Ralph Cicerone in 1986 to separate forcings (they called them thermal trapping) from responses (temperature changes), because the former can be done more accurately than the latter, and this allows a simple comparison of forcing agents. And in 1990 the first calculations of forcing by sulphate aerosols were presented. The latest publications by the Intergovernmental Panel on Climate Change illustrate the usefulness of this concept, showing comparisons of the magnitudes of forcings and the uncertainties about how different components affect the atmosphere.

Finally, I would have liked to have seen clear and concise explanations of what the greenhouse effect really is, how it works and what the basis is for scientific confidence (or lack thereof) in the basic theory and in forecasts of the future. While both authors provide solid knowledge in a retrospective sense, the images they give are not scientifically integrated and seemed to stop around the 1970s. But, rather than criticize them, I believe this is strong evidence that those of us in today's climate research community have failed to teach our subject adequately to the public, let alone to our kindred academics. ■

Robert J. Charlson is in the Department of Meteorology, Arrhenius Laboratory, University of Stockholm, S10691 Stockholm, Sweden.

Planting the evidence

A Rum Affair: How Botany's 'Piltdown Man' was Unmasked

by Karl Sabbagh

Penguin: 1999. 224 pp. £16.99

Christopher M. Berry

Many scientific scandals result from incompetence, poor methodology or the unexpected behaviour of equipment, but why might a researcher deliberately falsify results? As the pressure to produce results and publish has given just about every academic a motive, perhaps cases from the past might cast scientific fraud in a purer light, when the battle was for perceived truth, and not for a job or the next grant.

In 1948 John Raven, classics fellow of King's College, Cambridge, and a knowledgeable amateur botanist, set off on an



Scene of the crime: the Hebridean island of Rhum, home to the whole allegedly fraudulent affair.

undercover mission to the Scottish island of Rhum intent on proving that J. W. Heslop Harrison, the eminent naturalist and fellow of the Royal Society, was engaged in scientific fraud. Karl Sabbagh investigates the circumstances surrounding Raven's crusade using the personal accounts of witnesses, the increasingly acidic correspondence between the protagonists and the report Raven submitted to his college. Raven's accusation was that records of sedges and other plants recorded from Rhum and other Hebridean Islands, either rare or reported for the first time from Britain, were falsified by Heslop Harrison.

Raven's report has lain unread in the library of King's ever since. Sabbagh produces enough evidence, mainly from Raven but also from other sources, to show that Heslop Harrison did have some pretty damning charges to answer, both on the subject of the plants and also insect discoveries.

How does this scandal compare with other alleged acts of specimen-based scientific fraud? It appears that a small, rival group of British Hebridean botanists, more closely connected to the botanical establishment, were most affected by the Rhum affair, and that Raven was operating on their behalf.

Sabbagh sometimes seems to have difficulty persuading himself that the whole thing really mattered very much (British floral lists have been quietly corrected since). It had none of the public impact of Piltdown, and did not affect numbers of international specialists like the case of Talent vs Gupta. The fun of Piltdown is that it is a genuine 'who dunnit?', with a multitude of illustrious suspects ranging from museum technicians

and French priests to famous authors. *A Rum Affair* has but one suspect, so interest lies mainly in the discussion of potential motive.

Sabbagh suggests that Heslop Harrison was guided by his belief that some Hebridean islands were left unscathed by the last ice age, and that therefore the living biota should contain relics of the original flora and fauna as well as have similarities to southerly regions beyond the reach of the ice sheet. Frustrated in his attempts to have his theory recognized, he might have resorted to (literally) planting evidence so that others might appreciate the 'truth'. Rivalry and one-upmanship may also have contributed. But, of course, personality traits are as important as motive in such a situation, as the scientist has to step beyond the bounds of good practice. Sabbagh does well to construct portraits of the private and scientific personae of his two (deceased) leading men, perhaps telling us more than what is strictly on the page.

In *Unravelling Piltdown* (Random House, 1996), John Evangelist Walsh turns us against Charles Dawson, the 'discoverer' of Piltdown Man, with tales of shady property deals, sightings of sea monsters and dodgy archaeological artefacts. Sabbagh nails Heslop Harrison by discovering that he did not live on the rather grand and desirable street The Avenue, in Birtley, Yorkshire, as his letter headings claimed, but in the neighbouring and much less pretentious Ruskin Road. How such vanity and bending of the truth in his personal life (other examples are provided) might have affected his scientific work is a valid question. Yet details of his career, which might put him in a more favourable light, are missing. Why, for example, was

Heslop Harrison elected a Fellow of the Royal Society?

A *Rum Affair* is very entertaining; it is carefully written for the non-specialist, with many interesting diversions (such as Heslop Harrison's attempts to demonstrate Lamarckian evolution). It becomes uncomfortable (almost voyeuristic) reading as the ageing suspect is cornered and desperate to defend himself. Its appeal will doubtless spread beyond those interested in the natural history of Rhum and history of British botany to those fascinated by the wider issues of scientific fraud and those who enjoy a gripping yarn.

Christopher M. Berry is in the Department of Earth Sciences, Cardiff University, Cardiff CF1 3YE, UK.

Confessions of a modern roué

Nazis, Women and Molecular Biology: Memoirs of a Lucky Self-Hater

by Gunther S. Stent
Briones Books: 1998. 388 pp. \$25

Walter Gratzer

Günter Siegmund Stensch was born into a prosperous Jewish family in Berlin during the turbulent days of the Weimar Republic. In later life he cast off the heroic Wagnerian encumbrance of names his parents had wished on him, and as Gunther Stent became a phage geneticist of note and latterly a trenchant commentator on the ways of science. Now in his seventies, he has written this captivating memoir of his early life, to rid himself, perhaps, of the ghosts of a past that has seemingly continued to trouble him. For Stent is clearly a man corroded with guilt — the guilt of an “anti-Semitic Jew”, who tried, like so many of the fervently patriotic, assimilated German Jewish community, to hide his origins, to the extent indeed that he aspired to join the Hitler Youth; guilt, also, at the self-absorbed insouciance that pervaded his personal relationships.

Stent's recollections focus much more on his amorous than his scientific conquests. The young Stent was, by his own confession, a predatory Lothario, who repaid devotion with indifference. He has, it seems, preserved the letters, full of unrequited yearning, from the women with whom he shared his life while he was a member of the Technical Industrial Intelligence Branch; this was a rather futile body set up by the US Department of Commerce to monitor the state of German science and technology in the immediate aftermath of the Second World War. Stent's paramilitary standing allowed him to make a poignant visit to the

New in paperback

The Raptor and the Lamb: Predators and Prey in the Living World

by Christopher McGowan

Penguin, £8.99

“Christopher McGowan clearly loves animals and their adaptations, and this is a thoroughly enjoyable popular account, written for the reader who likes something meatier than a TV wildlife spectacular ... McGowan is at his best with the sheer physicality of hunting: the magnificence of top predators such as falcons, lions, cheetahs and hunting dogs, or the effects of viscosity on motion at the scale of the plankton. In its objectivity and refusal to be emotional, the text becomes very moving — for example, in the account of ‘suicide’ by an African buffalo cornered by five lions.” John R. G. Turner, *Nature* 396, 130–131 (1998)

Cosmology and Controversy: The Historical Development of Two Theories of the Universe

by Helge Kragh

Princeton University Press, \$19.95, £11.95

“The book recounts in often riveting detail how scientific interest in the theory of cosmology was awakened by Einstein nearly 80 years ago and how almost simultaneously the observation of the redshifts of distant galaxies by Slipher ushered in much wider involvement in the subject.”

Hermann Bondi, *Nature* 384, 323–324 (1996)

Figments of Reality: The Evolution of the Curious Mind

by Ian Stewart and Jack Cohen

Cambridge University Press, £9.95

“Figments will not appeal to everyone. It will hold few charms for everyone homozygous for the *humourless* mutagene, or without much imagination. But science without humour and imagination seems rather pointless — just like books without pictures or conversation were to

Alice. We need imagination to frame hypotheses, to wonder what is around the next corner, to ask “What if ...?”. And anyone without a sense of humour should be a merchant banker, not a scientist — the pay is better, for a start.”

Henry Gee, *Nature* 389, 452–453 (1997)

The Heavens on Fire: The Great Leonid Meteor Storms

by Mark Littmann

Cambridge University Press, £15.95, \$19.95

“Littmann tells his story with real flair, and expansively enough to teach a great deal of meteor, comet and meteorite astronomy in the process. There is considerable technical detail, mostly in side-bars, so the book is altogether satisfying both as a quick, informative read and also as a reference source.” Owen Gingerich, *Nature* 397, 33 (1999)

Nature Wars: People vs. Pests

by Mark L. Winston

Harvard University Press, \$15.95, £9.95

“In an articulate and accessible writing style, Winston explains the pesticide dilemma, the threat that our reliance on synthetic pesticides poses to both human health and safety and to the preservation of what is left of the natural environment ... Winston's discussion of these controversial issues ... will be helpful to anyone who hopes to develop an informed opinion about our continuing war with nature.” Lawrence M. Hanks, *Nature* 390, 573 (1997)

The Search for the Giant Squid

by Richard Ellis

Penguin, \$14.95

“A gold-mine of fact and fantasy, for we scientists who work on cephalopods and for all of us who love monsters.” Martin Wells, *Nature* 396, 641–642 (1998)

scenes of his childhood. It led also to his encounter with, by his own remorse-laden account, a warm, loyal and selfless young woman, whom, having made pregnant, he deserted.

Stent intersperses his narrative with flashbacks to episodes in his childhood — the suicide of his mother, his early exposure to anti-Semitic malice and his adventures in a militaristic Jewish youth organization of fascist leanings, disbanded by the Gestapo (which favoured instead the Zionist groups that promised to rid the Fatherland of its Jews). When schooling for Jewish children in German schools was proscribed, his father found a place for Stent in a Jewish boarding-school, which offered a liberal education with instruction in languages, in anticipation of an imminent dispersal of the children to France, England, the United States and Palestine. Nearly all escaped, enough of them

to America for a grand postwar reunion in New York with several of their teachers. Stent's father, brother and sister got away to England, leaving Stent himself and his new stepmother to join them after a hazardous last-minute exit by way of Holland. He did not linger long in London, but went on to Chicago, where his beloved sister and her husband were already settled.

The war and his tour of paramilitary duty in Germany over, Stent returned to the United States to complete his PhD in physical chemistry. But his studies gave him little satisfaction and he soon began to look for a more fulfilling career. He found what he was after in the austere but magnetic personality of another German emigré, Max Delbrück, the Pope of the new genetics. Delbrück at first rebuffed Stent's overtures, but relented when his young admirer reappeared with the highly coveted Merck fellowship in his pocket.

Stent paints a nostalgic picture of science and comradeship in Delbrück's research group in the postwar years, and at the Cold Spring Harbor Laboratory, where the geneticists met each summer and where Stent embarked on yet another ill-starred romance.

He does not dwell much on his work in phage genetics, which was then at the forefront of genetic research. He became a prominent figure in the field, although in the words of one of his friends, "Gunter is almost always wrong, but he is always interesting" (as he was in his widely read book, *The Coming of the Golden Age*, in which he asserted that molecular biology had passed its zenith and was now entering a state of stasis and tranquil satiety). In this, as in much else, he was a true follower of his adored patron Delbrück, who seems to have been repeatedly wrong in his scientific judgements, but whose rigorous and ascetic scepticism exerted an influence on genetics that led him eventually to a Nobel prize (which he shared with Salvador Luria and, according to Stent, accepted only after a struggle with his conscience).

Stent owns that he was never driven by any lofty urge to lay bare the secrets of nature. He enjoyed science for what Robert Oppenheimer called "the life it brings" — for the companionship and the give-and-take of debate that can be such a felicitous part of daily life in the laboratory. Were he alone on a desert island, he tells us, he would not bother with research. He is an astute observer of human foibles, and his concluding chapters are rich in anecdotes and vignettes of the great, the merely famous and the infamous; and over all hovers the brooding shade of Saint Max.

The most remarkable feature of Stent's book, aside from an engaging style, a pleasant wit, an amiable line in self-deprecation and an acute sense of history, is his extraordinary candour about his love affairs, his moral failures and the prosaic motives behind his science. Why has he so bared his soul to public view? I can only surmise that this is his way of exorcizing demons that have gnawed at his conscience since the days of his profligate youth. Stent introduces himself as a widower in his seventies and he draws a veil over his life after 1950, when his narrative ends. One must hope that he eventually found in his family life the serenity of mind that had for so long eluded him. ■

Walter Gratzer is at the Randall Institute, King's College London, University of London, 26–29 Drury Lane, London WC2B 5RL, UK.

Other memoirs

The Angry Genie: One Man's Walk through the Nuclear Age

by Karl Z. Morgan and Ken M. Peterson
University of Oklahoma Press, \$24.95

Science in culture



Animal arts

Martin Kemp

As soon as ancient philosophers attempted to differentiate human from animal intelligence, a few celebrated cases of 'animal geometers' attracted avid attention. The webs of spiders, the honeycombs of bees and the spiral shells of molluscs are the classic examples, to stand alongside the six-cornered snowflake and phyllotaxis. They testified to what Johannes Kepler called "Nature's formative virtue", the "*facultas formatrix*", through which God insinuated "world-building figures" into matter.

The building bee is eulogized in Kepler's *The Six-Cornered Snowflake* (1611): "the architecture is such that any cell shares not only six walls with the six cells in the same row, but also three plane surfaces on the base with three other cells from the contrary row". He compares this symmetrical packing to the seeds in a pomegranate, and attributes it to the same necessity as operates when pellets are systematically compressed in a round vessel.

The outline of the basic solution to the geometry of the walls was proposed by the fourth-century Greek mathematician Pappus of Alexandria. The preface to Book V of his *Collection* is devoted to "the Sagacity of Bees", to introduce "a problem of wider extent, namely that, in all equilateral and equiangular plane figures having an equal perimeter, that which has the greater number of angles is always greater". While bees are not accorded the reason needed to formulate such a general case, God has "granted that each of them should, by virtue of a certain natural instinct, obtain just so much as is needful to support life". Thus the bees "would necessarily think that the figures must be ... contiguous with each other ... in order that no foreign matter could enter ... and defile the purity of their produce".

Pappus knew that only three rectilinear, equiangular figures would fill the space: the triangle, the square and the hexagon. For their part, the bees, "by reason of their instinctive wisdom chose ... the figure which has the most angles because they conceived it would contain more honey".

But despite sustained attention from Kepler's successors, the full proof of why the hexagon delivers the maximum area for the minimum perimeter, compared to all other possible combinations of packed figures, remained elusive. The bees seemed to know more about the

isoperimetric problem than did mathematicians, including Sir Christopher Wren. The conundrum still warranted a substantial historical review by D'Arcy Wentworth Thompson in *On Growth and Form* in 1917 (Cambridge Univ. Press, 1992). Most recently, a proof of mighty dimensions, occupying 19 pages on the web, has been offered by Thomas Hales of the University of Michigan, an expert in geometrical packing (see www.math.lsa.umich.edu/~hales).

The effort needed to emulate the humble bee, brings us back to the question of animal intelligence with even greater force. To express the problem in modern terms, are the architectural bees obeying a genetic predisposition, or is the regularity the outcome of the self-organizing principles of packing? The wonderfully precise assembly of the thin cell walls at 60°-angles from tiny pellets of wax, to minute tolerances of 0.002 mm, excludes self-organization as the immediate mechanism. However, geometrical packing, like that of bubbles in foam, is as much a 'cause' of the configuration as genetic predisposition.

We can now see how Kepler's "formative virtue" elided two distinct but complementary processes. One is the kind of physical self-organization he observed in snowflakes, while the other is the instinctual programme that permits the bees to carry out such symmetrical acts of waxy engineering within the physical parameters of their world. It seems to me that the relative weights accorded to either of the interlocked processes in each complex instance of animal and vegetable artistry need to be argued case-by-case, even if the roles of geometry and genes now seem to be susceptible to clear definition in the long-running case of the honeycomb.

It is hardly surprising that such wondrous masterpieces of design have inspired artists and architects, no less than mathematicians. For Susan Derges, whose work is illustrated here — and can be seen in the book *Susan Derges, Liquid Form* (Michael Hue-Williams, 1999) — the honeycomb not only embodies the principles of natural order, but also encodes patterns of thought as the bees weave the tapestry of their industrious motions across the geometrical network. ■

Martin Kemp is in the Department of the History of Art, University of Oxford, 35 Beaumont Street, Oxford OX1 2PG, UK.

Spark ignites physicists

Hertz's work on electromagnetism started as many arguments as it settled.

Dominique Pestre

Physics textbooks tell us that Heinrich Hertz discovered electromagnetic waves in 1888. In so doing, Hertz proved experimentally that it was James Clerk Maxwell's theory of electromagnetic radiation, and not the idea of electrical action at a distance, defended by continental physicists such as Hermann Helmholtz, that was right. While basically correct, this story also contains elements which are largely false.

The first to seize upon Hertz's publications were, of course, the British 'Maxwellians', as they were already convinced that Maxwell was right. They welcomed the news that the physics professor at Karlsruhe had found ways to produce electromagnetic waves, to have them interfere, and to measure their speed of propagation in air, which he found to be the predicted 300,000 km per second. They announced the result everywhere and began mounting public demonstrations of the marvel of the sparks induced at a distance by a Hertzian generator.

But it was J. J. Thomson who did the first quantitative experiments. He was intrigued by Hertz's claim that the speed of propagation in wires was 200,000 km per second, two-thirds of that in air. Thomson conceived a more refined experiment in wires, declared the speed of propagation to be the same as that in air, and said he had no idea why Hertz had been mistaken. Two physicists from Geneva, Edouard Sarazin and Lucien de la Rive, found a more disquieting result at the end of 1889. After conducting numerous experiments they claimed that, contrary to what happens in optics, the distance between nodes and loops depended crucially on the size of the detector used to make the spark.

A professor of physics at the École Polytechnique in Paris, Alfred Cornu, used this finding to claim that Hertz's interpretation of his experiments was in trouble. Hertz's main hypothesis was that the generator produced a wave whose period T was unique and which he obtained by calculation. Hertz measured the distance L between two loops, from which he deduced the speed of propagation V . If L depended on the detector, Cornu concluded, T was not unique or V was variable. Either was problematic.

In the following months many physicists looked into these issues. Studies of the various parameters of this wonderful electrical phenomenon were undertaken, detectors and generators sprang up all over Europe, new setups were conceived, and diverse interpretations were put forward as solutions to

the problem. Some people were unmoved. The British Maxwellians thought the problem was trivial. If the detector was tuned to the generator before starting the experiment (something Hertz had been careful to do, and which he considered a necessary condition), he was simply right — and the question of the other wavelengths was not essential and could be solved later. Hertz had got to the crux of the matter, and the Swiss physicists had found a marginal phenomenon.

Another French savant, the mathematician Henri Poincaré, identified another problem. While he thought that the approximations Hertz needed in his calculations were acceptable, he noticed that Hertz had made a mistake in calculating the period of his generator, and that he was wrong by a factor of $\sqrt{2}$. This meant that if Hertz had actually measured the length L between two loops, the speed of propagation was 300,000 km per second multiplied by $\sqrt{2}$. Poincaré did not make a lot of this, in contrast to what would probably happen today. Accepting the extraordinary complexity of the experiment, nobody

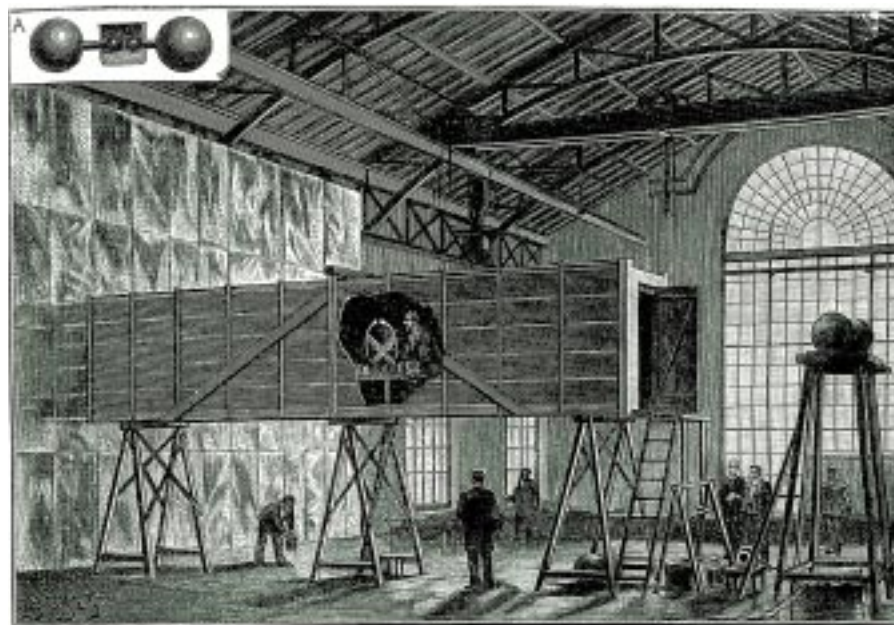
hinted at forgery or fraud, nobody thought (as far as we can judge) that Hertz's results were too good to be true. Poincaré simply concluded that he was willing to reconsider the whole quantitative question with the help of his close friend and experimental colleague, René-Prospère Blondlot. As for Sarazin and de la Rive's results, he suggested that they were because of the damping of the wave.

At no point did anybody doubt the decisiveness of Hertz's achievements. Many, many physicists entered the new domain and quickly performed experiments inspired by Hertz. All were able to generate sparks and 'play' with them, and all considered Hertz a true genius. On the other hand, most people were finally convinced by their own experiments, their own devices and calculations, their own way of adjusting proofs and expectations — and rarely by other people's (and in particular Hertz's) precise claims.

A fascinating indication of this is to be found in the textbooks published in the following two years by Hertz, Thomson and Poincaré. Each had his own story of what counted as a proof, and of what was decisive in securing the agreement of the community, and they differed profoundly. Hertz had made a major discovery, no doubt, but what he had proved, and who had decisively improved our understanding of this complex phenomenon, remained a matter of opinion.

Dominique Pestre is at the Centre d'Histoire des Sciences et des Techniques, École des Hautes Études en Sciences Sociales, Paris, France.

Nobody hinted at forgery or fraud, nobody thought that Hertz's results were too good to be true.



Switched on: Sarazin and de la Rive try to replicate and extend Hertz's experiments with electricity.

Life on the road

John G. Flanagan

As nerve axons migrate during wiring of the nervous system, molecular signposts at intermediate targets show them the way. Such a target has now been found to keep axons alive — provided they're on the right track.

To form the connections of the nervous system, axons must find their targets by setting out on a journey that may be long and complicated. How do they select the right path? What happens to those that stray off in the wrong direction? Intermediate targets are known to act as signposts along the way, sending out guidance molecules that tell the axons which way to go. On page 765 of this issue, Wang and Tessier-Lavigne¹ show that intermediate targets may have another function — they can also support the survival of migrating axons. Such a mechanism, by keeping axons alive only if they follow the right route, could ensure that the correct neural connections are formed.

The concept that nerves derive support from the targets they ultimately innervate can be traced back more than a century². The best-known way in which this support — now known as a neurotrophic effect — occurs is by prevention of programmed cell death. The idea is that, during development, an excess of neurons is produced initially. Their axons then compete for a limiting amount of neurotrophic factors in the target. The result is a match in the sizes of the nerve and target, as well as elimination of mis-targeted axons. A molecular basis for this phenomenon has been established through studies on nerve growth factor, the related neurotrophins, and other molecules that act as neurotrophic factors in final targets³.

Wang and Tessier-Lavigne¹ set out to investigate whether similar principles might apply to an intermediate target — an idea they refer to as *en passant* (in passing) neurotrophic activity. As their model system they chose spinal commissural axons (those that cross from one side of the spinal cord to the other), and their intermediate target was a wedge-shaped group of cells called the spinal-cord floor plate⁴ (Fig. 1a). Commissural axons are guided towards the floor plate, at least in part, by diffusible chemoattractants belonging to the netrin family. After crossing the midline, these axons turn and grow alongside the floor plate towards the brain, eventually heading off to reach their final targets.

As a first step, Wang and Tessier-Lavigne developed an *in vitro* assay. In the presence of netrin-1, rat dorsal spinal-cord explants send out commissural axons. But within days the axons degenerate and the cell bodies die. Could this degeneration be rescued by

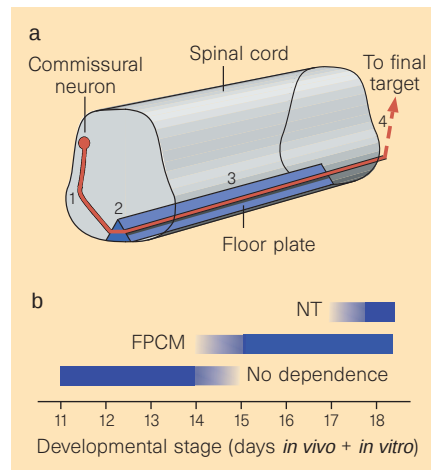


Figure 1 Commissural neurons and their intermediate target, the floor plate.

a, Commissural neurons in the rat dorsal spinal cord send out their axons, which soon turn (1) towards the floor plate, guided by chemoattractants. They then cross the floor plate (2) and turn again to grow alongside it (3) in the direction of the brain. Later they leave (4) to reach their final targets. b, Wang and Tessier-Lavigne¹ grew commissural axons from spinal-cord explants *in vitro*. They found that the explants degenerated at about the time they should normally have reached the floor plate. However, floor plate-conditioned culture medium (FPCM) contains a neurotrophic activity that can rescue them. Later, FPCM is not sufficient, but the axons can be rescued by FPCM if molecules in the neurotrophin family (NT) are also added.

the intermediate target? The authors found that the axon degeneration and cell death were indeed dramatically rescued with floor plate-conditioned culture medium (FPCM), revealing a powerful neurotrophic activity.

Is the activity localized to the intermediate target? The molecule responsible has not yet been identified, and, although it seems to be a polypeptide, tests of many candidates failed to identify any with appropriate activity. So, Wang and Tessier-Lavigne tested the localization by assaying slices taken from different parts of the spinal cord. They found that the activity is highly concentrated in the floor plate, with some also in more dorsal regions. These results indicate either lower production of the activity in dorsal regions, or diffusion of a signal produced in the floor plate.

If this activity is to fit the *en passant* neurotrophic model, a key issue is whether it can promote survival of cell bodies by acting at a distance on their axons. This was shown for nerve growth factor in classic experiments using the 'Campanot chamber', where axons were separated from their cell bodies by growth through a sealed barrier. Spinal commissural axons grow too slowly for this approach to be used, but Wang and Tessier-Lavigne developed an innovative and persuasive alternative. They placed a dorsal spinal-cord explant next to a floor plate explant in a collagen gel. By examining the axons and staining for dying cell bodies, they showed that, independent of the location of the cell bodies, the cells survived only if the axons had grown near the floor plate. This result is consistent with the idea that survival of commissural neurons in the embryo could depend on the distant journey of their axons.

So spatial aspects seem to fit the model — but what about time? *In vivo*, commissural axons grow out between embryonic day 11 (E11) and E13, and take about a day to reach the floor plate. Wang and Tessier-Lavigne found that, in the absence of FPCM, all of the dorsal spinal-cord explants looked healthy at the equivalent of E14. Yet, just one day later, all had begun to degenerate (Fig. 1b). This was true irrespective of whether the explants were placed in culture at E11 or E13, suggesting that a clock runs in the dorsal spinal cord *in vivo* or *in vitro*. This clock sets off a requirement for neurotrophic support just as the commissural axons are supposed to be reaching the floor plate. Axons that don't get there by the time the clock strikes presumably suffer a fate worse than Cinderella's.

Later, after the developing axons grow away from the floor plate, one might expect them to come to depend on neurotrophic activity from their final targets. Consistent with this idea, Wang and Tessier-Lavigne found that, at the equivalent of E18, axons can no longer be supported by FPCM. Instead, they need a combination of FPCM and neurotrophins (Fig. 1b). This could fit with a two-step model of development, in which the growing axons first depend on an *en passant* signal from the intermediate target, and later need an additional signal from the final target.

The *en passant* neurotrophic activity identified in these elegant studies could help solve the neural wiring problem in several



100 YEARS AGO

A bicycle ride will be none the less enjoyable if you train yourself, not merely to travel far, but to take an interest in the sights and scenes through which you pass. For the sake of example, let me remind you that no country is so rich as England in the architecture of its village churches. It is no hard matter to learn to recognise the principal peculiarities of the architectural types which prevailed from the days of the Saxons to Sir Christopher Wren. ... But as soon as the elements of English church architecture are known, an old church ceases to be merely a picturesque object. It is an historical document which you yourself can read. You do not need the aid of the sexton to tell you which is the oldest part. You can make a good guess at when that aisle was added, or that window knocked in a wall obviously older than itself. A visit to a cathedral becomes an intellectual pleasure. Weariness at the drone of the verger as he recites his oft-repeated lesson is replaced by an alert desire to know if the authorities from whom he learnt it confirm or correct the rapid conclusions as to date or history to which you yourself have come.

From *Nature* 19 October 1899.

50 YEARS AGO

In a recent investigation, I have even demonstrated, among other matters, the existence of red blood corpuscle remnants in ancient Swedish skeletons (Viking age), buried without the embalming procedures used in Egypt and elsewhere. Pictures of relatively well-preserved organic framework of bone tissue were obtained with material even from so early a period as the Upper Stone Age. Various substances have been identified in mummified tissues by means of chemical methods. ... It was now thought possible to ascertain ... whether histamine occurs in measurable amounts in mummy tissue and other ancient human remains. Soft tissue (skin from the neck) and bone (cervical vertebra) from a mummy of the Egyptian Museum in Stockholm and supposed to be about three thousand years old were ground to a fine powder. ... A definite spasmogenic activity was found in the extract thus prepared. The spasm of the isolated guinea pig's small intestine elicited with the extract could be prevented with an antihistamine drug and was to a certain degree counteracted by atropine.

From *Nature* 22 October 1949.

ways. First, if axons go astray, this mechanism might help to eliminate them quickly, before they interfere with the orderly pioneering and assortment of axon tracts. Second, it could prevent axons reaching the wrong final target, where they might otherwise be incorporated in aberrant neural circuits¹. Third, it opens the possibility of a combinatorial mechanism, where a limited number of factors derived from the intermediate or final targets could be used in different combinations to specify many distinct connections.

Wang and Tessier-Lavigne's work adds a fascinating new dimension to the increasing recognition that neurotrophic effects may go beyond the simple model of support by final targets. During development, the responsiveness of some neurons to different neurotrophins switches. Although this has not been tied unequivocally to intermediate targets, neurotrophins may contribute support at cell bodies or along pathways, while axons

are still on the way to their targets⁵. There is also some analogy in later events, when motor neurons require support both from their muscle target and from the Schwann cells that wrap around their axons after reaching the target³. We also have several new questions. What is the precise developmental significance of the floor plate activity detected *in vitro*? Could the work have therapeutic implications in spinal-cord regeneration? What molecules are responsible for the activity? We don't have all the answers yet — but after all, life's a journey. ■

John G. Flanagan is in the Department of Cell Biology and Program in Neuroscience, Harvard Medical School, Boston, Massachusetts 02115, USA. e-mail: flanagan@hms.harvard.edu

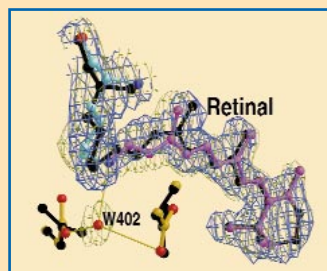
1. Wang, H. & Tessier-Lavigne, M. *Nature* **401**, 765–769 (1999).
2. Jacobson, M. *Developmental Neurobiology* (Plenum, New York, 1991).
3. Pettmann, B. & Henderson, C. E. *Neuron* **20**, 633–647 (1998).
4. Van Vactor, D. & Flanagan, J. G. *Neuron* **22**, 649–652 (1999).
5. Davies, A. M. *Curr. Opin. Neurobiol.* **7**, 110–118 (1997).

Protein crystallography

Frozen in time

Biochemical reactions are extremely rapid, but the methods for imaging the enzymes that catalyse them can take hours. To get round this problem, structures can be determined at temperatures around 100 K, literally freezing an enzyme's movements and allowing the intermediates in its reaction cycle to be observed. This has been done for bacteriorhodopsin in studies reported by Edman *et al.* in this issue (*Nature* 401, 822–826; 1999) and Genick *et al.* (*Science* 286, 255–260; 1999).

Bacteriorhodopsin is a pump that uses light energy to drive protons across bacterial purple membranes. A cycle of structural changes is triggered when absorption of a photon of light causes isomerization of bacteriorhodopsin's bound chromophore, retinal (purple in the figure). Intermediates in this photocycle can accumulate in crystals at



low temperatures (see *Nature* 392, 206–209; 1998) or in mutant proteins, as used by Genick *et al.*, but subtler techniques are needed to trap the earliest and most elusive intermediates.

Edman *et al.* maintained bacteriorhodopsin crystals in the dark, bathed in liquid nitrogen (110 K). They then drove the initial step in the protein's photocycle by illuminating the crystals with green light through an optical fibre (at a wavelength of 532 nm), and exposed them to a powerful synchrotron source to aid the rapid collection of data. The authors found that, despite the isomerization of a double bond, there is

very little change in the overall shape of the retinal — in the figure, electron density after illumination (blue) is compared with that before (brown).

But there are other changes. The biggest of these is the escape of a water molecule (designated W402) from the vicinity of the retinal. This molecule previously formed part of a network of water and amino-acid residues connecting retinal to the outside of the bacterial membrane. Its loss triggers changes in this network, eventually resulting in expulsion of a proton from the bacterial cell.

One snap-shot cannot reveal the whole picture of how the bacteriorhodopsin pump works. However, by building up series of freeze-frame structures for all the intermediates in the enzyme's working cycle the mechanics of this, and other, molecular machines is being revealed.

Christopher Surridge

Chemical physics

Molecules are cool

John M. Doyle and Bretislav Friedrich

Under ordinary conditions, atoms and molecules of a gas zigzag in all directions and have a wide distribution of speeds related to temperature. At room temperature, for instance, gaseous atoms or molecules are most likely to move at the speed of rifle bullets. Such restlessness puts a limit on the level of detail at which atomic and molecular properties can be studied.

The emergence of methods for slowing and trapping gaseous species has led to a renaissance in atomic physics, which is now progressing into molecular/chemical physics as well. The latest developments come from Bethlem *et al.*¹ and Maddi *et al.*² — they have independently devised kindred techniques for slowing molecules that potentially provide new approaches to subsequent trapping and spectroscopy in particular.

Control of atomic behaviour has already become pretty sophisticated. For instance, a class of atoms, best represented by the alkali metals, can now be routinely slowed with light and loaded into traps (constructed from light or magnetic fields)³. This 'laser cooling' relies on the rapid absorption and emission of photons by an atom placed in light tuned near the atom's resonant frequency. With a careful arrangement of the laser beams, the atoms can preferentially absorb photons from a single direction. Due to the momentum of the photon, the atom gets small velocity 'kicks' opposite to its direction of travel. This can be used to slow a beam of atoms so they can be caught in traps and cooled to very low temperatures (for example, evaporative cooling is applied to magnetically trapped atoms to attain Bose–Einstein condensation at temperatures below 1 μK). Unfortunately, the energy-level structure of molecules prevents the maintenance of the requisite simple absorption–emission cycle and so makes laser slowing impractical. In consequence, molecules (and many complex atoms) were relegated to the sidelines.

Last year, however, two very different approaches to cooling and trapping of molecules were successfully implemented. In one, CaH molecules were cooled with a cryogenically refrigerated helium buffer gas and loaded into a magnetic trap⁴. In the other, Cs₂ molecules were formed from laser-cooled Cs atoms and then confined in a light trap⁵. Now we have the techniques devised by Bethlem *et al.*¹ and Maddi *et al.*². Like laser slowing of a beam, they produce velocity 'kicks' opposite to the direction of travel. They work with bunches of atoms or molecules all travelling in one direction, and make

no recourse to lasers or cryogenics. Instead these new methods use time-varying inhomogeneous electric fields to provide the 'kicks' and the resultant slowing.

Inhomogeneous electric fields that can be turned on or off quickly are produced by pairs of electrodes. The potential energy of an atom or a molecule varies in a characteristic way with the strength of the electric field. Depending on whether the electric dipoles are on average antiparallel or parallel to the electric field, the particle's energy either increases (these are low-field seekers) or decreases (these are high-field seekers) with increasing field strength. The energy of low-field seekers is lowest at minimum field strength and, therefore, in an inhomogeneous field they seek these low-field regions. In contrast, high-field seekers are forced towards the maximum field strength.

Polar molecules possess both high- and

low-field-seeking states in an electric field. Atoms, on the other hand, become polarized only in the presence of an electric field and so their electric dipoles are always parallel to the field; they can only be high-field seekers in an electric field. The technique of Bethlem *et al.* relies on low-field seekers and, therefore, is well suited for slowing polar molecules; that of Maddi *et al.* relies on high-field-seeking states and so is potentially suitable for slowing both atoms and molecules.

Figure 1 illustrates the two approaches. In the scheme of Bethlem *et al.* (Fig. 1a), molecules enter an inhomogeneous electric field from a field-free region. As they approach the region of maximum field strength between the electrodes, the potential energy of low-field seekers is increasing. This occurs at the expense of their translational energy which is being correspondingly reduced (a consequence of energy conservation). To prevent the molecules from regaining the translational energy as they leave the field, the field is quickly switched off. The molecules thus end up in a homogeneous, field-free region where their potential (and thus translational) energy remains constant. To maximize the loss of translational energy, the field has to be

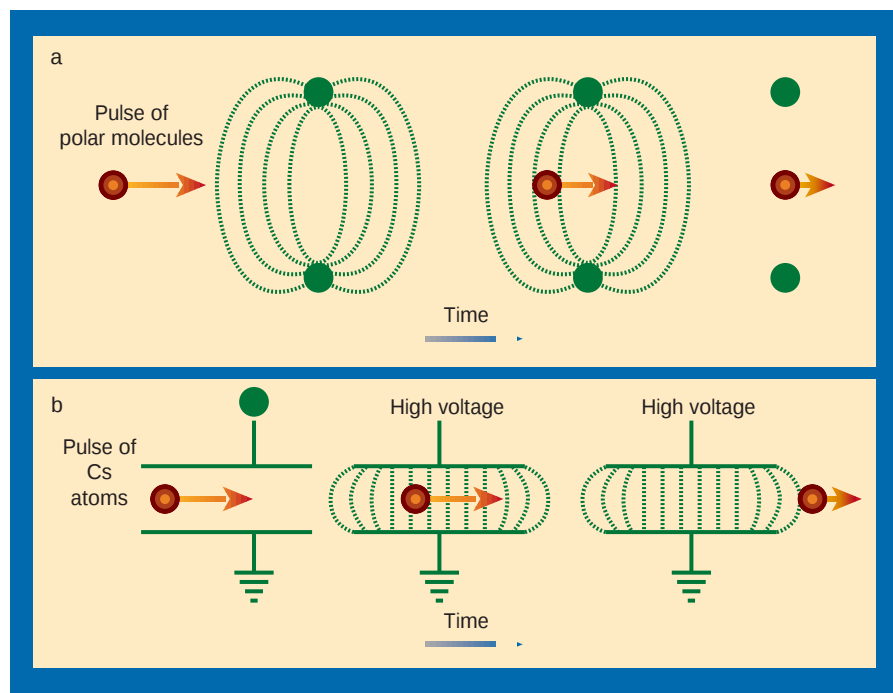


Figure 1 Principle of molecular and atomic slowing by time-varying inhomogeneous fields. Length of orange arrows indicates speed. a, One of the 63 stages used by Bethlem *et al.*¹ to decelerate a pulsed beam of CO($a^3\Pi$) molecules from 225 m s^{-1} to 98 m s^{-1} . The two opposing rods are simultaneously switched by two independent high-voltage switches. The decelerating stages are stacked in alternating vertical and horizontal configurations, keeping the molecules focused. They gain potential energy as they approach the field maximum, and so lose kinetic energy. When the field is switched off, the molecules are left with reduced velocity. b, A deceleration stage used by Maddi *et al.*² to slow a pulse of Cs atoms (released from a magneto-optic trap) from 2 m s^{-1} to 0.2 m s^{-1} . The two opposing condenser plates are simultaneously energized by two independent high-voltage power supplies. The electric field is turned on when the Cs pulse reaches the uniform field region, thereby inducing an electric dipole in the atoms. As the pulse leaves the plates it passes through an area of decreasing field strength. Its potential energy increases at the expense of kinetic energy, resulting in pulse deceleration.

switched off just at the point where the molecules reach the field's maximum.

Maddi *et al.* (Fig. 1b) used atoms in their experiments, but their scheme could, in principle, be applied to molecules. The atoms enter an inhomogeneous electric field from a region of a large homogeneous electric field. The homogeneous field is formed by two condenser electrodes that are energized after the atoms' entry. Atoms in their high-field-seeking states then gain potential energy as they pass through the inhomogeneous field at the exit of the condenser, again at the expense of their translational energy.

In both versions, the atoms or molecules arrive in pulses: a pulsed beam of metastable molecules is created from a supersonic beam expansion (Bethlem *et al.*), or a pulsed beam of ground-state atoms is launched from a magneto-optic trap (Maddi *et al.*). The pulsing is crucial, because the switching of the fields must be accurately synchronized with the arrival of the atoms or molecules. These techniques may enable atoms or molecules to be slowed enough for capture in traps.

Yet another new approach⁶ to slowing beams of atoms and molecules is based on supersonic expansion from a nozzle placed at the end of a spinning armature. The nozzle is oriented so that it moves in the opposite direction to the discharging gas, and its speed roughly equals the speed of the expanding atoms or molecules. As a result, it cancels the particles' speed in the laboratory frame, leaving the atoms or molecules essentially at rest.

Cold molecules should prove useful in spectroscopy and the study of molecular structure, especially in ultra-high-resolution spectroscopy, which requires cold (slow) and trapped (long-interaction-time) samples. An especially promising area for study is collisions of ultra-cold molecules, when the molecules behave like waves, perhaps giving rise to a new chemistry. The technique may also enable the study of collective quantum effects in molecular systems, including Bose–Einstein condensation. Just as atom cooling is opening up new avenues of research, it is likely that the same will happen with molecular cooling — with repercussions for chemistry and even, perhaps, biology.

John M. Doyle is in the Department of Physics, and Bretislav Friedrich is in the Department of Chemistry and Chemical Biology, and the Department of Physics, Harvard University, Cambridge, Massachusetts 02138, USA.
e-mails: jd@jsbach.harvard.edu
brich@chemistry.harvard.edu

- Bethlem, H. L., Berden, G. & Meijer, G. *Phys. Rev. Lett.* **83**, 1558–1561 (1999).
- Maddi, J. A., Dinneen, T. P. & Gould, H. *Phys. Rev. A* (in the press). <http://xxx.lanl.gov/abs/physics/9909027>
- Wiemann, C. E., Pritchard, D. E. & Wineland, D. J. *Rev. Mod. Phys.* **71**, S253–S262 (1999).
- Weinstein, J. D. *et al. Nature* **395**, 148–150 (1998).
- Takekoshi, T., Patterson, B. M. & Knize, R. J. *Phys. Rev. Lett.* **81**, 5105–5108 (1998).
- Herschbach, D. *Rev. Mod. Phys.* **71**, S411–S418 (1999).

Pharmacology

Towards better benzodiazepines

William Wisden and David N. Stephens

Sleeplessness and anxiety afflict us all — sometimes to the point of breakdown. The usual treatment is drugs of the benzodiazepine family, such as diazepam (Valium), which have helped millions of patients since they were introduced in the late 1950s. Benzodiazepines induce relaxation, but they can also play tricks on memory (hence their association with date rape), reduce concentration, cause physical clumsiness and intensify the effects of alcohol. They can also be addictive. In an exciting study on page 796 of this issue, Rudolph *et al.*¹ report the use of mouse molecular genetics to study the target for benzodiazepine action. This work sets the foundations for developing Valium-like drugs that ease anxiety or help the onset of sleep, but without the destructive side effects.

To function properly, brain circuits need balanced positive and negative signals. The negative inhibitory signals come from the

chemical neurotransmitter GABA (γ -aminobutyric acid). When GABA binds to its target protein complex — the GABA_A receptor — a chloride ion-channel opens, allowing chloride ions to move into neurons and damp down their electrical activity. Benzodiazepines bind to a specific site on the GABA_A receptor and induce a subtle change in its shape. This change causes GABA to work more efficiently at the receptor, so the movement of chloride ions into the neuron increases^{2,3}. But because GABA_A receptors are found throughout the brain, benzodiazepine treatments often affect those circuits that control, say, attention or balance, as strongly as they influence the circuits that regulate anxiety; hence the side effects.

The situation is further complicated because there are many types of GABA_A receptor^{2,3}. Most are made of α , β and γ subunits arranged in a pentamer, with the chloride ion-channel at the centre. Benzodiazepines bind at the interface between the α and γ subunits⁴ (Fig. 1). There are six types of α subunit, $\alpha 1$ – $\alpha 6$, distributed in various brain circuits to make an overlapping mosaic of receptor subtypes. So it is difficult to understand what each subtype does. Available drugs are not selective enough, and genetic studies in mice — where the gene for a particular subunit is inactivated — often cause death⁵ or severe neurological illness⁶. This is because the brain circuitry has no brakes on its activity, so it spirals into dysfunction.

In an alternative approach, Rudolph *et al.*¹ have exploited a variable amino acid in the α subunit known as the R/H site⁷. The $\alpha 1$, $\alpha 2$, $\alpha 3$ and $\alpha 5$ subunits all have histidine (H) at position 101 of their polypeptide chain, whereas $\alpha 4$ and $\alpha 6$ have arginine (R)⁷. The histidine contributes to the binding pocket for benzodiazepines, and *in vitro* studies⁷ show that only GABA_A receptors with this amino acid at position 101 of their α subunits are sensitive to these drugs. If histidine is mutated to arginine, drugs such as diazepam no longer improve the action of GABA at the GABA_A receptor. However, the receptor's sensitivity to GABA and its function stay unchanged⁸.

Rudolph *et al.* now put this discovery into a physiological setting. Using the technique of homologous recombination, they engineered mice in which residue 101 in the $\alpha 1$ subunit is mutated from histidine to arginine. (The $\alpha 1$ protein is a good first choice for this experiment because it contributes to over 60% of GABA_A receptors in the brain.) The authors found that drug-free

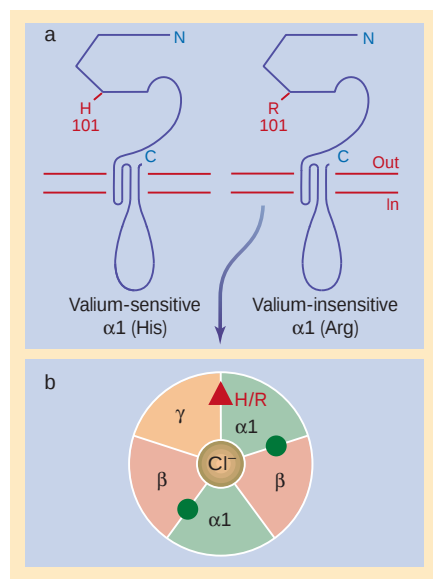


Figure 1 Structure of the GABA_A receptor. The receptor is made up of five subunits, and one of these, the $\alpha 1$ subunit, is shown from the side (a). In a complete receptor, the subunits form a ring with the chloride (Cl^-) ion-channel at the centre (shown from above, b). The neurotransmitter GABA binds at the interface between the α and β subunits (green circles), causing the chloride channel to open. The interface between an $\alpha 1$ and a γ subunit creates the binding site (red triangle) for benzodiazepine drugs such as Valium. If a key histidine residue (H101) in the extracellular amino-terminal domain of the $\alpha 1$ subunit is changed to arginine, the benzodiazepine-binding site is inactivated although the receptor still responds to GABA^{4,7}.

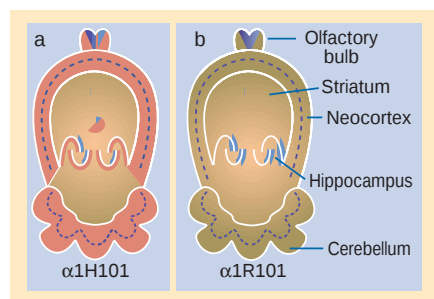


Figure 2 Distribution of benzodiazepine-binding sites in the mouse brain. a, In brain with a normal $\alpha 1$ gene ($\alpha 1H101$), the location of benzodiazepine-sensitive GABA_A receptors with an $\alpha 1$ subunit is shown in pink, and GABA_A receptors containing the $\alpha 2$, $\alpha 3$ and $\alpha 5$ subunits are in blue. Some parts of the brain (such as the hippocampus, cortex, olfactory bulb and thalamus) contain a mixture of receptor types, indicated by the overlap of blue and pink. Other brain areas, such as the striatum, contain mainly non- $\alpha 1$ receptors. b, In the $\alpha 1R101$ mice made by Rudolph *et al.*¹, the $\alpha 1$ -containing receptors cannot bind benzodiazepines and only the $\alpha 2$, $\alpha 3$ and $\alpha 5$ sites remain. Large areas of the $\alpha 1R101$ brain are insensitive — or, at least, less sensitive — to benzodiazepines.

$\alpha 1(H101R)$ mice are healthy, ending speculation about the existence of natural benzodiazepine-like modulators. If histidine 101 were part of the binding site for these hypothetical molecules, this site would have vanished in the $\alpha 1(H101R)$ mice, so they might, for example, have become more alert or slept less. But Rudolph *et al.* found no hints of such behaviour.

When the authors gave the mutant mice diazepam, however, there were two clear changes. First, the mice showed no sedation (measured by how much they run around and show an interest in their environment). And second, their memory was not impaired — at least, not for unpleasant events (going into a dark chamber and receiving an electric footshock). In contrast, when non-mutant mice were given the same dose of the drug they became sluggish, relaxed and apt to forget bad experiences.

These experiments highlight the importance of $\alpha 1$ -containing GABA_A receptors in the brain circuits that control general arousal and certain types of memory. However, the anxiety-reducing properties of benzodiazepines are just as potent in $\alpha 1(H101R)$ mice as in controls. In other words, the anxiety-reducing effects of benzodiazepines are distinct from their general sedative actions (until now, an open question), and must depend instead on GABA_A receptors with $\alpha 2$, $\alpha 3$ or $\alpha 5$ subunits.

Maintenance of muscle tone also seems to be independent of $\alpha 1$ -containing receptors, because the $\alpha 1(H101R)$ mutants became clumsy when treated with diazepam. And the ability of benzodiazepine drugs to enhance

the effects of alcohol — a fact reflected in road-death statistics — is also independent of the $\alpha 1$ subunit, as the $\alpha 1(H101R)$ mice were affected just as strongly as wild-type animals by alcohol/benzodiazepine cocktails. Receptors containing the $\alpha 2$ and $\alpha 3$ subunits (which, unlike those containing $\alpha 1$, are expressed in the spinal cord) are likely candidates. But $\alpha 1$ subunits might still be involved in the other ill effects of combining alcohol and benzodiazepines, such as memory blackouts.

Can drugs be developed that are specific for GABA_A-receptor subtypes, and hence for, say, anxiety-reducing effects? 'Proof-of-principle' successes with $\alpha 5$ - and $\alpha 6$ -selective compounds^{9,10} indicate that such drugs are not far away. The technique described by Rudolph *et al.* is equally applicable to genes for the $\alpha 2$, $\alpha 3$ and $\alpha 5$ subunits (Fig. 2). The corresponding H101R mutants will narrow down which behaviours are influenced by which receptor type, and these results can be combined with large-scale screening for drugs on cloned receptors. The promise for

the future is a directed search for drugs to treat conditions including anxiety, insomnia and epilepsy.

William Wisden is at the MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK.

e-mail: ww1@mrc-lmb.cam.ac.uk

David N. Stephens is in the Department of Experimental Psychology, University of Sussex, Falmer, Brighton BN1 9QG, UK.

e-mail: dns@biols.susx.ac.uk

1. Rudolph, U. *et al.* *Nature* **401**, 796–800 (1999).
2. Barnard, E. A. *et al.* *Pharmacol. Rev.* **50**, 291–313 (1998).
3. Mckernan, R. M. & Whiting, P. *Trends Neurosci.* **19**, 139–143 (1996).
4. Sigel, E. & Buhr, A. *Trends Pharmacol. Sci.* **18**, 425–429 (1997).
5. Günther, U. *et al.* *Proc. Natl Acad. Sci. USA* **92**, 7749–7753 (1995).
6. Homanics, G. E. *et al.* *Proc. Natl Acad. Sci. USA* **94**, 4143–4148 (1997).
7. Wieland, H. A., Lüddens, H. & Seeburg, P. H. *J. Biol. Chem.* **267**, 1426–1429 (1992).
8. Kleingor, C., Wieland, H. A., Korpi, E. R., Seeburg, P. H. & Kettenmann, H. *Neuroreport* **4**, 187–190 (1993).
9. Korpi, E. R. & Lüddens, H. *Br. J. Pharmacol.* **120**, 741–748 (1997).
10. Sur, C., Fresu, L., Howell, O., Mckernan, R. M. & Atack, J. R. *Brain Res.* **822**, 265–270 (1999).

Carbon cycle

The blast in the past

Gerald R. Dickens

On current estimates^{1–3}, over a period of less than a thousand years 2,000–4,000 gigatonnes of carbon will be added to the atmosphere by human activity. That's 2–4 billion billion tonnes. What will be the consequence of this rapid and massive release of carbon? The question has been tackled primarily with numerical simulations of the global carbon cycle constrained by experiments, present-day observations and records from the late Quaternary, the past 200,000 years or so of Earth history.

An alternative — studying ancient blasts of carbon — has always seemed pointless simply because we thought that there weren't any such blasts; as many of us know, natural processes cannot suddenly add enormous amounts of carbon to the ocean or atmosphere². That view of the global carbon cycle is spectacularly flawed, however, as highlighted in the paper by Norris and Röhl on page 775 of this issue⁴.

For just a brief period, about 55 million years ago, temperatures at high latitudes and in the deep oceans soared by 5–7 °C (refs 5–7). This event, called the late Palaeocene thermal maximum or LPTM⁶, coincided with an extraordinary decrease in the ¹³C/¹²C ratio ($\delta^{13}\text{C}$) of all carbon on the Earth's surface^{5,7–9}. In sequences of marine sediments examined to date, this isotope anomaly is an abrupt drop in $\delta^{13}\text{C}$ of –2.5 to –3 ‰ over 5 to 20 cm followed by a roughly exponential

return to near-initial values over 1 to 4 m (refs 5,7,9). The magnitude, shape and widespread nature of the isotope anomaly indicate a massive input of carbon to the ocean or atmosphere from an external reservoir greatly enriched in ¹²C, followed by a gradual return to earlier conditions^{7,9,10}.

Time is central to understanding carbon input during the LPTM. Based on long-term sedimentation rates at several marine locations, previous studies^{5,7,9} suggested that the abrupt drop and gradual return in $\delta^{13}\text{C}$ spanned less than 10,000 years and about 140,000 years, respectively. But such an interpretation necessarily implies that carbon was being added to the ocean or atmosphere at rates approaching or exceeding those of present-day anthropogenic inputs^{7,10}. Clearly, we have a dilemma: either models of the global carbon cycle are incomplete at a basic level, or the inferred timing of the $\delta^{13}\text{C}$ anomaly is incorrect. Without a detailed timescale across the LPTM, one can certainly attribute the apparently rapid occurrence of the $\delta^{13}\text{C}$ anomaly to a fanciful interpretation of the geological record.

During 1997, the Ocean Drilling Program recovered a continuous sediment sequence across the LPTM from a hole (site 1051) in the western Atlantic Ocean^{4,11}. Unlike all previously recovered late Palaeocene sections, this record contains obvious cycles in sediment composition.

What Norris and Röhl⁴ have done is to quantify the amplitude and thickness of each cycle over a considerable length of the upper Palaeocene and lower Eocene, between 54 and 56 million years ago. Then, following numerous previous studies of marine sediment, they show that the observed sediment cycles have a regular frequency consistent with the Earth's precessional period of about 20,000 years. Using this astronomically calibrated timescale for the Palaeocene–Eocene

transition, they conclude that the abrupt decrease and gradual return in $\delta^{13}\text{C}$ at site 1051 indeed took less than 10,000 years and about 140,000 years, respectively (Fig. 1). The full 2.5–3‰ amplitude of the isotope anomaly is not evident in the bulk carbonate record, probably because the bulk record includes variable amounts of carbon from deep- and shallow-ocean reservoirs. Nevertheless, the data concerned look solid: there is now no question that the isotope composi-

tion of all major carbon reservoirs suddenly changed during the LPTM.

These new records across the LPTM force us to reconstruct the widely held models of the global carbon cycle. Specifically, there must be a large reservoir that, in the distant past, has sporadically been able to add huge quantities of ^{12}C -rich carbon to the ocean or atmosphere over very short periods of time. As discussed by Norris and Röhl⁴, and others^{9–11}, the best explanation is that the carbon cycle contains a large gas-hydrate capacitor (Fig. 2) because no other reservoir contains a large mass of carbon sufficiently enriched in ^{12}C . Gas hydrates are ice-like solids composed of gas and water that are stable at high pressure, low temperature and high gas concentration¹². We know that sediment on continental slopes contains enormous quantities of ^{12}C -rich carbon as CH_4 (methane) hydrate¹², and that has presumably been the case throughout time. Carbon input fluxes to this global gas-hydrate capacitor (conversion of organic matter to CH_4) have probably been similar to carbon output fluxes (oxidation of CH_4 to ΣCO_2) over most of the geological record. But during the 5–7°C deep-ocean warming at the LPTM, the cause of which remains unclear, it seems that thermal dissociation of gas hydrate caused CH_4 release rates greatly to exceed production rates¹⁰. The result was a massive and rapid input of at least 1,000 gigatonnes of ^{12}C -rich carbon into the ocean and atmosphere (Figs 1, 2).

Carbon we add to the atmosphere today will eventually precipitate as carbonate and organic carbon^{2,3}. In the debate over the parameters and feedbacks that govern such carbon sequestering, present-day observations and Quaternary records offer only limited help. In theory, future removal of anthropogenic input of carbon into the environment could be tracked by a quasi-exponential increase in the $\delta^{13}\text{C}$ of the ocean or atmosphere^{3,13}. If the carbon cycle 55 million years ago operated in a similar way to the way it does now, the record presented by Norris and Röhl⁴ implies that the carbon we inject over the coming millennium will remain with us in the ocean, atmosphere and biomass for at least another 120,000 years.

During the LPTM, the first 10,000 years after massive carbon input when $\delta^{13}\text{C}$ is relatively constant are particularly important, because this indicates a transition period when inputs and outputs of the global carbon cycle are balanced. Norris and Röhl⁴ and Bains *et al.*¹¹ suggest that this was a time of diminished CH_4 input (Fig. 2). A global carbon cycle that cannot remove carbon immediately after massive carbon injection is a provocative alternative. With a revised model for the global carbon cycle, and a single, well-dated late Palaeocene section, we can now begin to view aspects of Earth's future in an entirely new light.

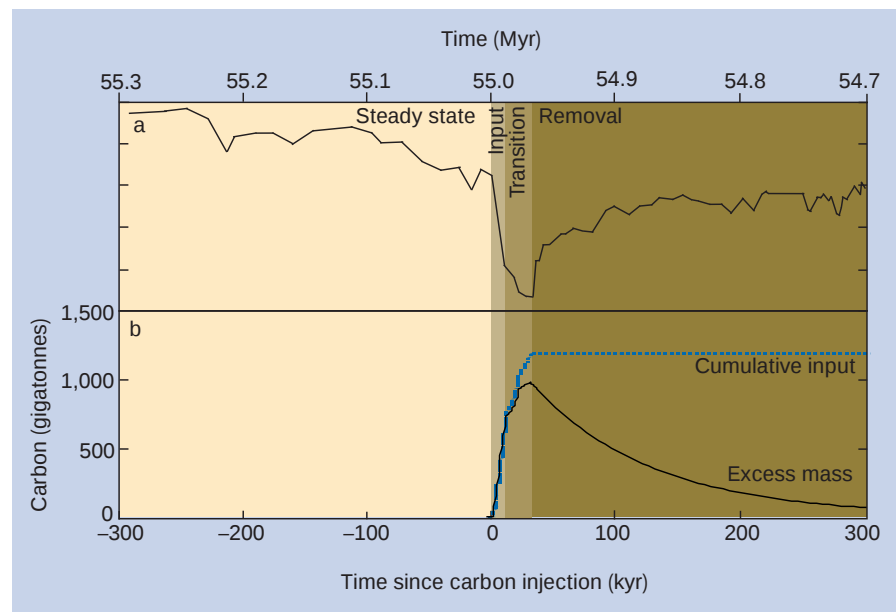


Figure 1 States of carbon around the late Palaeocene thermal maximum (LPTM), 55 million years ago. **a**, Norris and Röhl's astronomically calibrated carbon isotope ($\delta^{13}\text{C}$) record across the LPTM at Ocean Drilling Program site 1051. **b**, Step-wise integration of the time-record according to mass-balance equations^{10,13} to obtain a methane carbon-source function. This simple analysis of the $\delta^{13}\text{C}$ record gives a minimum carbon input during the LPTM in gigatonnes. It assumes present-day values for the global carbon cycle^{2,10} and does not consider additional carbon input from dissolution of carbonate.

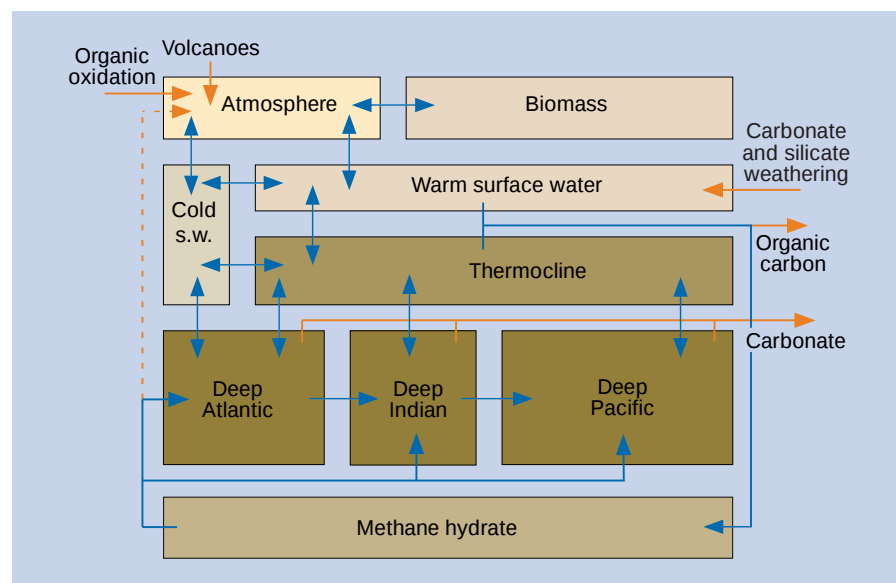


Figure 2 Model of the exogenic carbon cycle that includes a large gas-hydrate capacitor. The major carbon reservoirs, including gas hydrates, are connected by internal exchange fluxes (blue arrows). Carbon is added to and removed from these reservoirs by external exchange fluxes with the rock cycle (orange arrows). The sizes (masses) of reservoirs are not drawn to scale.

Gerald R. Dickens is in the School of Earth Sciences, James Cook University, Townsville, Queensland 4811, Australia.

e-mail: Jerry.Dickens@jcu.edu.au

- Carbon Dioxide Information Analysis Center (CDIAC) <http://cdiac.esd.ornl.gov/trends/trends.htm>
- Sundquist, E. T. in *The Changing Carbon Cycle: A Global Analysis* (eds Trabalka, J. R. & Reichle, D. E.) 371–402 (Springer, New York, 1986).
- Walker, J. C. G. & Kasting, J. F. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **97**, 151–189 (1992).
- Norris, R. D. & Röhl, U. *Nature* **401**, 775–778 (1999).
- Kennett, J. P. & Stott, L. D. *Nature* **353**, 225–229 (1991).
- Zachos, J. C., Lohmann, K. C., Walker, J. C. G. & Wise, S. W. *J. Geol.* **101**, 191–213 (1993).
- Thomas, E. & Shackleton, N. J. *Correlations of the Early Paleogene in Northwest Europe* Geol. Soc. Spec. Publ. 101 (eds Knox, R. O. et al.) 401–411 (Geological Society, London, 1996).
- Koch, P. L., Zachos, J. C. & Gingerich, P. D. *Nature* **358**, 319–322 (1992).
- Bralower, T. J. et al. *Geology* **25**, 963–966 (1997).
- Dickens, G. R., O'Neil, J. R., Rea, D. K. & Owen, R. M. *Paleoceanography* **10**, 965–971 (1995).
- Bains, S., Corfield, R. M. & Norris, R. D. *Science* **285**, 724–727 (1999).
- Kvenvolden, K. A. *Rev. Geophys.* **31**, 173–187 (1993).
- Berger, W. H. & Vincent, E. *Geol. Rundsch.* **75**, 249–269 (1986).

Immunology

Toll gates for pathogen selection

Richard J. Ulevitch

All species need an immediate, systemic reply to the microbial pathogens in their environment. This reply — known as the innate immune response — is characterized by the *de novo* production of mediators that either kill the pathogen directly or induce phagocytic cells to ingest and kill it. In the fruit fly *Drosophila melanogaster*, distinct cell-surface receptors belonging to the Toll family mediate separate anti-bacterial and anti-fungal responses in the same type of cell¹. They do this by inducing genes that encode anti-microbial peptides.

Does this model describe the situation in mammalian systems? On page 811 of this issue, Underhill *et al.*² describe the ligand specificity of two members of the mammalian Toll-like receptor family, TLR2 and TLR4. These data show that the situation is indeed similar to that in *Drosophila* — that is, two distinct Toll-like receptors, both expressed in macrophages, discriminate between different microbial pathogens or their products. Whereas TLR4 is the main protein involved in recognizing Gram-negative bacteria and lipopolysaccharide (a glycolipid constituent of the bacterial outer membrane), TLR2 is the key in responses to other types of microbial pathogen, such as yeast and Gram-positive bacteria.

The innate immune response is often driven through recognition, by the host cell, of surface components of the microbial pathogen. Monocytes and macrophages are central to the intensity and specificity of this response. Most — if not all — microbial pathogens are thought to activate the innate immune system using mechanisms with common features. For example, many macrophages contain a surface protein called CD14, which binds ligands such as lipopolysaccharide and triggers an innate immune response³. But CD14 is not thought to participate directly in signalling. Rather, one or more of the mammalian Toll-like

receptors acts in concert with CD14 to discriminate between microbial pathogens or their products¹ and initiate transmembrane signalling.

Mammalian cells may express as many as ten distinct Toll-like receptors⁴. All span the

cell membrane, with repeating leucine-rich repeats in the external domain and a common sequence motif — the Toll-homology domain — in the cytoplasmic tail. Members of the interleukin-1 receptor family, which is another essential element of the innate immune system, also contain a Toll-homology domain in their cytoplasmic tails. Most attention has been paid to the TLR2 and TLR4 proteins, and the importance of TLR4 in responses to Gram-negative bacteria and lipopolysaccharide was first suggested by powerful genetic data. Beutler and colleagues⁵ established that TLR4 is encoded by the lipopolysaccharide (*Lps*) gene, and that it controls sensitivity to this molecule. Akira and colleagues⁶ also showed that mice in which the *TLR4* gene has been deleted have the same defects as mice with mutations in *Lps*.

But some reports indicate that lipopolysaccharide acts via TLR2. When TLR2 is expressed in cell lines that do not normally produce it, these cells become able to respond to lipopolysaccharide. Moreover, in at least one case, lipopolysaccharide-induced activation of a monocytic cell line is

Earth science

Mercurial vents



On the land, many hot springs are enriched in trace metals, and have laid down economically important deposits of gold, silver and mercury. Now, Peter Stoffers from Kiel University and co-workers from Canada and New Zealand have discovered the first mercury-producing hot springs under the sea, in water 200 m deep off the coast of New Zealand's North Island (*Geology* **27**, 931–934; 1999).

The mercury-depositing vents are part of more than 20 in two

groups discovered in the Whakatane graben, a fault-bounded depression on the sea floor. They are inhospitable, if colourful, places, reaching temperatures of 200 °C, laying down thick crusts of arsenic and sulphur, and producing liquid hydrocarbons from the 'cracking' of organic matter — many of the samples came up coated in an oily film that smelt of petroleum. The vents are inhabited by mats of bacteria that can metabolize sulphur.

The vents are also almost saturated with mercury. Much of it is present as cinnabar (HgS; the red crusts in the picture), but Stoffers *et al.* estimate that about 10% exists as free mercury, some of which may have condensed from a mercury vapour formed inside the boiling vents. The water pressure forces some of the mercury into pores in

the rocks, silver droplets of which are released when the samples are brought to the surface.

The origin of the mercury is uncertain. The basement rocks beneath the vents are shales and graywackes (a form of sandstone). These rock types have been found to contain large mercury deposits which formed in the past and, as this result shows, formation may still be going on. Shales are rich in hydrocarbons, which would explain their presence in the vents. Mercury could also be entering the vents from nearby volcanoes. Stoffers *et al.* speculate that hydrothermal vents around the world could be adding sizeable amounts of mercury to the sea. That in some places this mercury could end up in fish and eventually humans is an obvious, although unquantified, concern.

John Whitfield

R. KELLY

blocked by an anti-TLR2 monoclonal antibody⁷. So TLR2 has been considered a lipopolysaccharide receptor. But do these findings accurately reflect the physiological pathways of innate immunity?

These uncertainties can now be put aside thanks to the data from Underhill *et al.*² and to a report by Takeuchi *et al.*⁸ in this month's *Immunity*. These authors found that mice lacking TLR2 respond to lipopolysaccharide in the same way as do wild-type animals. They characterized responsiveness to lipopolysaccharide in both whole animals and macrophages. When the authors looked at macrophages from their TLR2-deficient mice, they found them to be less responsive than those from wild-type mice to cell-wall preparations from several distinct Gram-positive bacteria. These data not only support the idea that TLR4 is the essential element of a lipopolysaccharide–receptor complex that determines subsequent cellular responses to lipopolysaccharide, but they also point to a key role for TLR2 in innate immune responses to other microbial pathogens and their products. And although the initial findings implicating TLR2 as a lipopolysaccharide receptor seem unlikely to be physiologically relevant, they should nonetheless be recognized for having opened a new chapter in innate immunity by showing that Toll-like receptors can activate cells in a ligand-specific way.

Where does this rapidly emerging field go now? Once again, studies in *Drosophila* may lead the way. For instance, Levashina *et al.*⁹ have highlighted the importance of proteases in controlling expression of the *spaetzle/Toll/cactus* genes in an anti-fungal defence system. This group showed that the absence of a serine protease inhibitor (serpin) encoded by the *Spn43Ac* gene results in constitutive activation of an innate immune response in the fly. Their findings support two main ideas: first, that the protease system activated by exposure to a microbial pathogen (or a product derived from that pathogen) is an essential control point in innate immunity; and second, that the Toll protein is not itself a pattern-recognition receptor in *Drosophila*¹⁰.

Given the parallels between the innate immune response in flies and in man, it is reasonable to look towards determining the ligand and specificity of all the Toll-like receptors. In mammals, are the pathogen-activated protease cascades represented by proteins in the complement and coagulation pathways? Or will careful analyses reveal that cell-surface protease systems work in concert with CD14 and members of the Toll-like-receptor family to generate innate immune responses? Whatever the case, such studies should uncover in more detail the conserved ancestral features of the two systems of innate immunity that are now being worked out in the *Drosophila* and mammalian systems. ■

Richard J. Ulevitch is in the Department of Immunology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037, USA.

e-mail: ulevitch@scripps.edu

1. Williams, M. J., Rodriguez, A., Kimbrell, D. A. & Eldon, E. D. *EMBO J.* **16**, 6120–6130 (1997).
2. Underhill, D. M. *et al.* *Nature* **401**, 811–815 (1999).
3. Ulevitch, R. J. & Tobias, P. S. *Annu. Rev. Immunol.* **13**, 437–457 (1995).

4. Rock, F. L., Hardiman, G., Timans, J. C., Kastelein, R. A. & Bazan, J. F. *Proc. Natl Acad. Sci. USA* **95**, 588–593 (1998).
5. Poltorak, A. *et al.* *Science* **282**, 2085–2088 (1998).
6. Hoshino, K. *et al.* *J. Immunol.* **12**, 3749–3752 (1999).
7. Brightbill, H. D. *et al.* *Science* **285**, 732–736 (1999).
8. Takeuchi, O. *et al.* *Immunity* **11**, 443–451 (1999).
9. Levashina, E. A. *et al.* *Science* **285**, 1917–1919 (1999).
10. Medzhitov, R. & Janeway, C. A. Jr *Proc. Natl Acad. Sci. USA* **95**, 429–430 (1998).

Fluid dynamics

Order in chaos

Hassan Aref

Vast quantities of fluid are mixed every day, both naturally and for industrial purposes, and one might think that mixing processes are fully understood scientifically. That is far from the case. New insights are continually emerging, as exemplified by the paper by Rothstein, Henry and Gollub on page 770 of this issue¹. They have experimentally tackled a type of low-speed fluid mixing known as chaotic advection², and describe the surprisingly stable spatial patterns that may arise.

Most mixing occurs in turbulent flows, where the flow field itself is composed of a complex hierarchy of interacting eddies of various sizes. These eddies will each mix fluid on their own scale. But what happens when turbulence does not occur? How does one mix fluids efficiently if the amount of agitation is somehow limited either in spatial extent or amplitude? This may occur, for example, when the fluid is very viscous or tightly confined.

For slow flows we speak of laminar as opposed to turbulent mixing, and the issue here is how to achieve efficient, laminar mixing. Everyday experience provides some ideas. Stirring milk or sugar in a cup of tea involves laminar flow. It is a task we understand intuitively. For an even slower flow consider the process of kneading dough, a mixing process for a very viscous substance. In this case, a different strategy is used in which the various layers of 'fluid' are bent and folded so as to bring initially distant parts close together and separate initially close parts.

This is our clue — laminar flows may be set up so that complex deformations of the fluid arise. In technical terms the deformation may constitute a chaotic mapping, in the sense of the theory of dynamical systems, between the initial configuration and a later configuration (for example the configuration one period later for a periodic agitation process). It turns out that the equations of motion for individual fluid particles may have chaotic solutions even if the flow field itself is a simple, laminar flow. This was

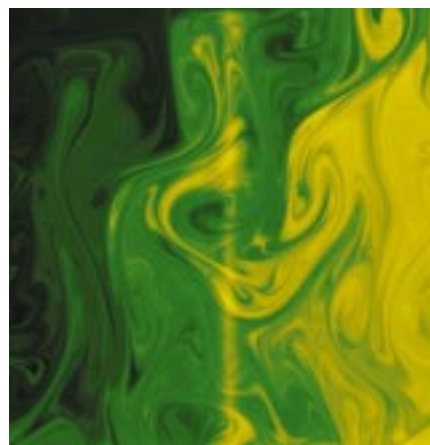


Figure 1 Recurrent pattern created by Rothstein *et al.*¹ using electromagnetic forces to stir a thin, viscous fluid layer. One side has been labelled by fluorescent dye. The periodic laminar flow exhibits chaotic advection within which a persistent pattern recurs once per cycle of the forcing. The pattern gradually fades out over about 100 cycles, like the grin on the Cheshire cat, while the striations maintain constant thickness.

shown by Arnold³ and Hénon⁴ in the mid-1960s, albeit for a very special class of flows (in fact, flows that do not contain any viscous effects at all). Their terse papers demonstrating how mixing in a laminar flow can display interspersed regions of regular and chaotic motion were, unfortunately, all but ignored by the science and engineering community.

The terms 'stirring' and 'mixing' enter the discussion, and require explanation. They might seem synonymous, but they have different technical meanings⁵. Stirring is the purely mechanical movement and rearrangement of fluid; mixing, on the other hand, occurs through diffusion — that is, through a homogenizing process at the molecular scale. A fluid system that simply homogenizes due to diffusion, then, is 'mixed but not stirred'. Conversely, if diffusion is essentially absent, agitation gives a fluid that is 'stirred but not mixed'. Chaos allows the creation of such fine striations,

folds and interleaving of material surfaces that by stirring alone we can approach arbitrarily closely to the fully mixed state. In particular, the spatial structure of the stirred fluid displays scales that are orders of magnitude smaller than what one would guess from an instantaneous snapshot of the streamlines of the flow.

So the motion — advection, as it is known — of individual particles in the fluid can be either regular ('integrable' in the terminology of dynamical systems) or chaotic. In two dimensions, chaotic advection² requires a time-dependent flow. In three dimensions it can occur in steady flow. Chaotic advection enhances the quality of mixing, because one of the hallmarks of chaos is the exponential separation of initially nearby points. If all pairs of points in some region of the flow separate exponentially, interfaces stretch dramatically and the end result for the bulk fluid is improved mixing. Over the past 15 years this idea of chaotic advection has become a mainstay of the liter-

ature on fluid stirring and mixing. The immediate applications to the mixing of viscous fluids in chemical engineering have been supplemented by more esoteric examples^{6,7}, such as mixing in physiological flows, in microfluidic channels, in geophysical fluids such as molten rock, and even in tidal flows in shallow waters.

Now Rothstein *et al.*¹ add an exciting chapter to this story. They demonstrate a major difference between chaotic, laminar mixing and turbulent mixing. It turns out that there is relative order in chaos: the complex spatial patterns created by chaotic advection due to periodic stirring of the fluid have a simple time dependence. Indeed, they re-emerge periodically with remarkable fidelity (Fig. 1).

Rothstein and colleagues describe an elegant experimental system for investigating these patterns. It consists of a thin fluid layer, confined to execute two-dimensional motion by density stratification and magnetic forcing, and stirred by passing a hori-

zontal electric current through the fluid in the presence of an array of fixed magnets. They then discuss various statistical measures of the patterns, comparing chaotic advection to advection by a weakly turbulent flow. In the latter case the complex patterns do not re-emerge. Regular (integrable) advection has simple spatial structure and time dependence. Chaotic advection produces complex spatial structure with simple time dependence. In turbulent advection there is complexity both in space and time.

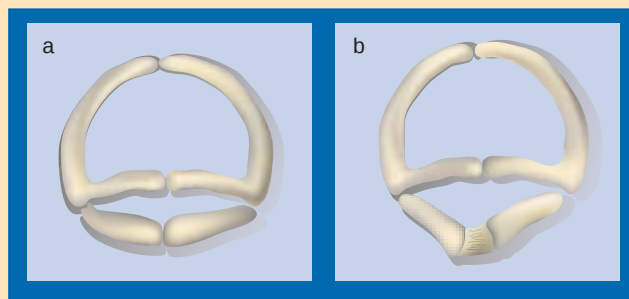
This phenomenon may have much broader significance. For instance, it is tempting to speculate that long-chain organic molecules sloshing about in a primordial soup could have been both well mixed and organized into patterns by chaotic advection. Indeed, the fine striations may have structure reaching down to molecular scales and, as Rothstein *et al.* show, simple time-periodic forcing will make these patterns re-emerge during every cycle. And to take another example, the rhythm of the pulsatile

Zoology

Lickety split

The giant anteater, *Myrmecophaga tridactyla*, possesses what may be the most specialized head morphology and feeding system among terrestrial mammals. With its elongate, tubular snout, its tiny, toothless mouth and its long, slender, sticky tongue, *M. tridactyla* can probe deep into the nests of ants and termites. Virginia Naples has presented a new analysis of how these remarkable creatures cope with their tricky diet (*J. Zool.* 249, 19–41; 1999).

Naples has discovered that the giant anteater has an unusual way of opening its mouth. Most mammals do this by contracting the digastric muscle to move the



lower jaw. The giant anteater has no digastric muscle, and the lower jaw depresses by only a few degrees during feeding. Instead, it opens its mouth by rotating the two halves of its elongate lower jaw (the mandibular rami) about their long axes.

Near the tip of the snout, the mandibular rami are flat, horizontal blades. They are connected only by a ligament, and so are able to rotate somewhat independently. To open its mouth, an anteater rotates the rami, depressing the inside edges of the blades and causing the flattened, oval mouth (a in the diagram above) to become a deeper diamond shape (b). This allows an anteater to open

and close its mouth rapidly, and permits feeding in ant and termite nests, where the spaces are too narrow for the mouth to open wide.

Naples proposes that rotation of the mandibular rami is produced by superficial portions of the masseter and temporalis muscles. This hypothesis is exciting and unexpected because in most mammals the masseter and temporalis act to close the jaw, not open it.

These findings suggest that, during anteater evolution, there has been a major change in the coordination of feeding muscles. In most mammals, the muscles that close the jaw contract just after tongue retraction, but Naples' model predicts that in giant anteaters

the masseter and temporalis muscles must be active during tongue protraction — don't try this at home, or you'll bite your tongue.

In recent years, functional morphologists have discovered that the motor coordination of muscles in different animal groups is surprisingly similar, despite large anatomical changes during evolution (for example, from tetrapod forelimbs to bird wings; see Dial, K. P., Goslow, G. E. & Jenkins, F. A. *J. Morphol.* 207, 327–344; 1991). The anteater may have been able to buck this trend because its jaws are not needed for chewing, freeing up jaw-closing muscles to be used for jaw rotating and mouth opening.

Naples' model relies mainly on anatomical evidence, and should be tested by measurements of muscle activity. If the model holds, it will stand as a rare example of a major evolutionary change in the motor coordination of muscle groups. **Elizabeth Brainerd**
Elizabeth Brainerd is in the Department of Biology, University of Massachusetts, Amherst, Massachusetts 01003-5810, USA. e-mail: brainerd@bio.umass.edu

flow in a cardiac system may serve both to increase mixing and to maintain organized spatial patterns.

Hassan Aref is in the Department of Theoretical and Applied Mechanics, University of Illinois at Urbana-Champaign, 104 South Wright Street, Urbana, Illinois 61801, USA.
e-mail: h-aref@uiuc.edu

1. Rothstein, D., Henry, E. & Gollub, J. P. *Nature* **401**, 770–772 (1999).
2. Aref, H. *J. Fluid Mech.* **143**, 1–21 (1984).
3. Arnold, V. I. *C. R. Acad. Sci. A* **261**, 17–20 (1965).
4. Hénon, M. *C. R. Acad. Sci. A* **262**, 312–314 (1966).
5. Eckart, C. *J. Mar. Res.* **7**, 265–275 (1948).
6. Ottino, J. M. *The Kinematics of Mixing* (Cambridge Univ. Press, 1989).
7. Aref, H. & El Naschie, M. S. (eds) *Chaos Applied to Fluid Mixing* (Pergamon, Oxford, 1995).

Space biology

Life without gravity

Richard J. Wassersug

Now is a quiet period in space biology, perhaps as quiet as before autumn 1957, when Laika the dog — and the discipline — left the launch pad. The primary platforms for biological research in microgravity are retired (Mir and NASA's modular Spacelab), on indefinite hold (Russia's unmanned BION satellites), or still being made (the International Space Station). Even NASA's space shuttles are grounded for maintenance. So this is a good point at which to take stock of where space life science has been and where it might go. As part of this process, NASA's Center for Advanced Studies in the Space Life Sciences sponsored an unusual workshop last month* on how microgravity affects biological systems — unusual in that it focused on microgravity's effects above the cellular level, in contrast to most contemporary biology, which is more molecular and reductionist in scope.

Research in space biology is often rationalized on the grounds that it may lead to medical advances for debilitating malaises of old age (such as osteoporosis and muscle wasting), or that it may improve the health and welfare of astronauts on long missions. Those arguments, although potentially valid, were hardly heard at the workshop. Rather, the common motivating force seemed to be a genuine desire to understand the role that gravity has played in the evolution of life on our planet, using microgravity as an investigative tool.

The workshop was unusual not only in its focus, but also in the diversity of organisms and organ systems discussed. Attendees heard about molluscs, insects, fish, amphibians, rats, monkeys, humans and even wheat^{1–4}. The word 'integrative' in the workshop's title refers to the direct and indirect effects of microgravity and other environmental variables on multiple organ systems and their development. A nice example of this is seen in medaka. This fish was the first

(and still the only) vertebrate to mate and lay eggs that developed into free-living offspring in microgravity⁵. Under these conditions, some strains of medaka continuously somersault as they swim. The problem is not in their vestibular system, as one might suppose, but in their visual system (K. Ijiri, Univ. Tokyo) — strains with poor vision are more likely to loop.

Another example is wheat. This is the first species to be raised to seed, from seed that was itself raised in microgravity (G. Bingham, Utah State Univ.). That simple goal — of getting a large terrestrial plant to complete two generations in space — had eluded space botanists for decades. We now know that water requires special management in microgravity. A granular matrix around roots, even a seemingly well-hydrated one, can have non-wet voids that hinder water uptake. And capillarity, unopposed by gravity, can produce films of fluid that block gas diffusion and limit oxygen uptake by roots. Similarly, in microgravity germinating seeds easily asphyxiate.

There is evidence from recent missions such as Neurolab that, during certain critical periods, gravity is necessary for normal development of the vestibular-motor reflexes in both frogs (E. Horn, Univ. Ulm) and rats (A. Ronca, NASA Ames Research Center, California). Developmental biologists at the workshop spoke for longer-term studies into this gravity dependence. Likewise, the comparative physiologists identified a multitude of questions that could be answered given longer-term missions. For instance, does the rate of bone demineralization in astronauts carry an increased risk of kidney-stone formation (C. Wade, NASA Ames Research Center, California)? Does the well-documented wasting of postural muscle in microgravity feed back on the nervous system such that reconditioning is retarded on re-entering 'normal' gravity (R. Roy, Univ. California, Los Angeles)? And does the unexpectedly high deposition of small (less than 1 µm) aerosol particles in the lungs under microgravity increase the risk of chest infections in astronauts (J. West, Univ. California, San Diego)?

Perhaps the most impressive demonstration of what can be accomplished came from neurobehavioural studies. On the 1998 Neurolab mission, four rats had multi-electrode recording arrays chronically implanted next to their hippocampal 'place cells' (neurons that encode a cognitive map of the environment), to see which cells fired as the animals negotiated a three-dimensional track (J. Knierim, Univ. Texas, Houston). Such studies could reveal how information on gravity (and other accelerations) is processed by the nervous system, from the peripheral receptor to deep within the brain. They could also help to explain the visual illusions experienced by some astronauts when they arrive in space — such as the impression that the world is upside down — and how the nervous system accommodates to microgravity as these illusions disappear later in the mission.

Despite a general consensus about what can realistically be accomplished in integrative space biology, there were subtle differences of opinion about what should fly on the International Space Station. For example, developmental biologists felt that research in the near term must use only a limited set of model organisms. In contrast, several physiologists and behaviourists felt that many species are needed, covering a large size range, so that scaling effects can be dissected out. Both are correct. For most studies of early embryogenesis, a rat is a rat. But for studies in physiology or behaviour, the differences between laboratory and wild strains of the same species may be enormous.

Although integrative space biology is barely 40 years old, the workshop showed that it has a pedigree of performance, as well as realistic goals. It was also agreed that the International Space Station should be an outstanding laboratory for integrative biology, despite its limitations for other scientific disciplines. To understand how microgravity affects biological systems we need the time, work space, equipment and on-site attention available only in a high-quality, manned space station with good internal environmental controls — just what the International Space Station is supposed to be.

Richard J. Wassersug is in the Department of Anatomy and Neurobiology, Dalhousie University, Halifax, Nova Scotia, Canada B3H 4H7.
e-mail: tadpole@is.dal.ca

1. Souza, K., Hogan, R. & Ballard, R. (eds) *Life into Space: Space Life Sciences Experiments, NASA Ames Research Center 1965–1990* (Government Printing Office, Moffett Field, California, 1995).
2. Moore, D., Bie, P. & Oser, H. (eds) *Biological and Medical Research in Space: An Overview of Life Sciences Research in Microgravity* (Springer, Germany, 1996).
3. Stark, R. E. *Ethology in Space, a Unique Opportunity for Behavioural Science* (ESA, Noordwijk, The Netherlands, 1993).
4. Nicogossian, A. E., Huntoon, C. L. & Pool, S. L. (eds) *Space Physiology and Medicine* (Lea & Febiger, Philadelphia, 1994).
5. Ijiri, I. *The First Vertebrate Mating in Space: A Fish Story* (RICUT, Tokyo, 1995).

* *Microgravity's Effects on Biological Systems and Behavior: An Integrative Approach*. NASA Center for Advanced Studies in the Space Life Sciences at the Marine Biological Laboratory, Woods Hole, Massachusetts, 2–4 September 1999.

HIV

See a pocket, block it

John P. Moore and Tatjana Dragic

Although drugs such as protease and reverse-transcriptase inhibitors have had a huge effect on treating AIDS in the developed world, they are not without their problems: they do not eradicate HIV-1 from infected people¹; long-term side effects have been observed²; and drug-resistant variants of HIV-1 are emerging³. There is a pressing need to identify new classes of anti-HIV-1 drugs, and, reporting in *Cell*, Eckert *et al.*⁴ signal a way forward. They have discovered that a cavity formed in the HIV-1 gp41 transmembrane glycoprotein can accommodate small, circular D-peptides — peptides that could prevent HIV-1 from fusing with, and entering, human cells.

Central to the immunodeficiencies of AIDS is a block to production of T cells bearing the cell-surface protein called CD4. When drugs reduce viral replication, the immune system gains time to repair this block⁵. The gp41 trimer mediates the initial fusion of HIV-1, and its potential as a drug target was revealed when large peptides derived from gp41 were found^{6,7} to inhibit HIV-1 infection *in vitro*. These peptides insert a metaphorical spanner into the works — they inhibit conformational changes within gp41, substituting for one or more components of the trimer and preventing the necessary intermolecular interactions that drive membrane fusion^{8,9}.

One version of these peptides, known as T-20, has been shown in clinical trials to reduce viral load, an important proof of concept for HIV-1 fusion inhibitors¹⁰. But large peptides make imperfect drugs because they cannot be administered orally and they are rapidly metabolized. Indeed, most successful drugs have a relative molecular mass of less than 1,000. So do any inhibitors of gp41-mediated fusion not have the drawbacks of conventional peptides? Eckert and colleagues' results⁴ indicate that some, to a certain extent at least, do.

The same group previously described^{11,12} a cavity in gp41 that seemed to be of a suitable size and location to act as the binding site for a fusion inhibitor. This cavity is not identical to the binding site for T-20 and related peptides, but it is a part of it. Eckert *et al.* have now engineered a simplified version of the cavity, designated IQN17, by fusing a soluble, trimeric helical peptide (GCN4-pI_{QI}) to 17 residues of the relevant gp41 region. This, in effect, produces a stable gp41 intermediate⁴. The authors then used IQN17 as a bait in a sophisticated screening procedure, based on phage-display library technology, to fish out the D-peptides that bind

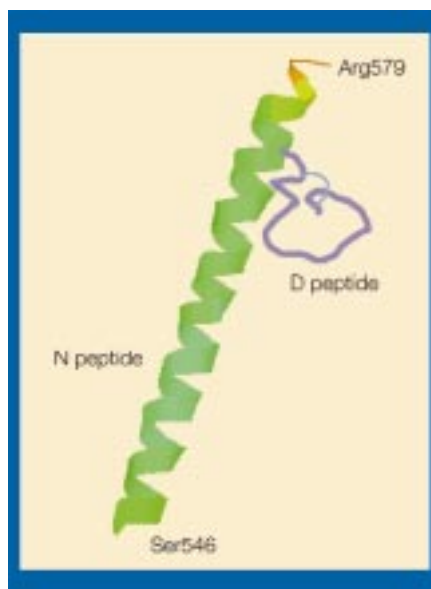


Figure 1 Plugging the gp41 pocket. Ribbon representation of a circular D-peptide inhibitor (purple) bound to an α -helical stretch of gp41 (green). Of the 16 residues in the peptide, only six interact directly with the gp41 pocket.

from a random pool. They showed that these peptides bind within the pocket of authentic gp41, successfully inhibiting membrane fusion (Fig. 1).

Why did the authors choose D-peptides? These peptides contain D-amino acids, whereas all natural peptides contain the mirror-image L-form. But, because D-amino acids are unnatural, the peptide bonds between them resist digestion by cellular enzymes, increasing the stability of the peptide *in vivo*. Indeed, Eckert *et al.* note that

cyclosporine A, a widely used drug that can be given orally to counter transplant rejection, contains D-amino acids and is not much smaller than the best of their inhibitory D-peptides.

The current D-peptide inhibitors are not very potent, with IC_{50} values (the concentration of inhibitor that is needed to block gp41-mediated fusion by half) in the 10–100- μ M range. However, the skilled hands of medicinal chemists can work wonders with starting compounds. Furthermore, Eckert and colleagues outline a plausible strategy in which IQN17 and the currently available D-peptides could be used in high-throughput screening assays to identify small-molecule inhibitors from conventional chemical libraries. There are many steps along the path from the bench to the pharmacy, but the first is often the most difficult. Together with inhibitors aimed at other components of the HIV-1 fusion process¹³, derivatives of the newly identified D-peptides may eventually be valuable weapons with which to fight HIV-1 infection. ■

John P. Moore and Tatjana Dragic are at the Aaron Diamond AIDS Research Center, The Rockefeller University, New York, New York 10016, USA.

e-mail: jmoore@adarc.org

1. Finzi, D. *et al.* *Nature Med.* **5**, 512–517 (1999).
2. Lucas, G. M., Chaisson, R. E. & Moore, R. D. *Ann. Intern. Med.* **131**, 81–87 (1999).
3. Yerly, S. *et al.* *Lancet* **354**, 729–733 (1999).
4. Eckert, D. M., Malashkevich, V. N., Hong, L. H., Carr, P. A. & Kim, P. S. *Cell* **99**, 103–115 (1999).
5. Rowland-Jones, S. *Lancet* **354**, 5–7 (1999).
6. Wild, C., Oas, T., McDaniel, C., Bolognesi, D. & Matthews, T. *Proc. Natl Acad. Sci. USA* **89**, 10537–10541 (1992).
7. Jiang, S., Lin, K., Strick, N. & Neurath, A. R. *Nature* **365**, 113 (1993).
8. Furuta, R. A., Wild, C. T., Weng, Y. & Weiss, C. D. *Nature Struct. Biol.* **5**, 276–279 (1998).
9. Jones, P. L., Korte, T. & Blumenthal, R. J. *Biol. Chem.* **273**, 404–409 (1998).
10. Kilby, J. M. *et al.* *Nature Med.* **4**, 1302–1307 (1998).
11. Chan, D. C., Fass, D., Berger, J. M. & Kim, P. S. *Cell* **89**, 263–273 (1997).
12. Chan, D. C., Chutkowski, C. T. & Kim, P. S. *Proc. Natl Acad. Sci. USA* **95**, 15613–15617 (1998).
13. Michael, N. L. & Moore, J. P. *Nature Med.* **5**, 740–741 (1999).

Computational neuroscience

Think positive to find parts

Bartlett W. Mel

How do you tell a shovel from an acorn? A bicycle from a mango? A young face from an old one? They look different, of course, but what exactly are the 'features' or 'parts' that our brains use to tell one thing from another? On page 788 of this issue, Lee and Seung¹ present a new algorithm for learning object parts. Known as non-negative matrix factorization (NMF), their algorithm follows from the simple prescription that parts are collections of input elements that tend to come and go together,

and which can be added together in different combinations to give objects of interest.

In the visual realm, the authors' view of parts is similar to the idea behind identikit — the collections of mix-and-match face parts used to reconstruct the identities of wanted fugitives or lost children. Parts that are too much like whole objects are rejected, because this would require an impracticably large inventory. Lee and Seung's approach also shuns parts which, although few in number, must be combined in counter-

intuitive mixtures to reconstruct objects of interest.

Learning techniques (the core of 'neural network' theory) have been used in various ways to find components of objects or images^{2,3}. Lee and Seung take a so-called unsupervised approach, the goal of which is to find structures — or 'basis functions' — that help to 'explain' a training set, in the sense that training examples can be faithfully reconstructed using appropriate combinations of the discovered basis functions. What distinguishes Lee and Seung's approach from other structure-finding algorithms is the way in which the basis functions must be combined to build target objects — these authors say that the combinations should be exclusively additive. Their approach reflects the idea that objects are most naturally described by the inventory of the parts they contain. Or, to look at things the opposite way, parts are precisely those entities that allow objects to be reassembled using purely additive combinations.

When Lee and Seung asked their NMF algorithm to decompose a set of faces, they found that the resulting basis functions were visibly different — and more part-like — than those found by conventional structure-finding algorithms such as principal components analysis (PCA) or vector quantization (VQ), although see ref. 4. Unlike NMF, PCA decomposes images into conceptually opaque additive and subtractive combinations of basis functions. The VQ technique represents objects using a set of whole-object prototypes, similar to the way in which suitcases are identified at an airline's lost-baggage counter. In other words, PCA finds parts that are not very intuitive and VQ ignores part structure altogether.

What is the importance of Lee and Seung's contribution? Most compelling is the argument that NMF's non-negativity constraint is well matched to our intuitive ideas about decomposition into parts. Second is the fact that the resulting basis functions differ in an interesting way from those generated without the non-negativity constraint. And third is the elegance and simplicity of the algorithm.

But there are caveats too. The authors rely on subjective appraisal — the look and feel — of their results to support the claim that NMF is better at finding the underlying component-based structure of complex objects than, say, PCA. And although Lee and Seung find that the NMF basis functions learned from faces seem to consist of "several versions of mouths, noses and other facial parts", another observer might describe many of these same basis functions as partial renderings of several face parts in unintuitive combinations (see Fig. 1 on page 789). Moreover, a simple threshold applied to PCA or VQ basis functions could considerably reduce the differences between these more

'global' basis functions and the more 'local' NMF ones.

Another question is how NMF basis functions relate to 'parts' as conventionally defined⁵. Given that the basis functions depend only on coactivation of input elements, a blob in a basis function extracted from a set of facial images could arise from the collusion of nose pixels in one image, upper-lip pixels in another and shadow pixels in a third. That is, unrelated entities, or 'causes', could be lumped within the same so-called NMF 'part'. But this display of poor partsmanship is less a commentary on NMF than it is a consequence of the input format (raw pixel values) and simplified representational scheme (weighted sums). Together these variables make it difficult for NMF, or any comparable scheme, to discover the deeper semantic structure that is normally associated with a parts-based representation.

Lee and Seung draw part of the inspiration for their non-negativity constraint from biology — that is, they point out that the firing rates of neurons are never negative, and that synaptic strengths do not change sign. But is this germane? Receptive fields in the brain often come in opponent pairs, where the positive and negative ranges of a variable are represented by the positive firing rates of two separate neurons (for example, 'on'- and 'off'-centre cells). Moreover, individual neurons, which are presumed to be the feature detectors in our own visual systems, typically show one or more kinds of opponency within their receptive fields. So a stimulus that is excitatory in one part of a receptive field may be inhibitory in another. In short, positive and negative quantities seem to be intermixed throughout our perceptual systems.

Despite the caveats, Lee and Seung's work is clear and thought-provoking, bearing on a scientific question of great importance — how to extract meaningful patterns from a complex world. To boot, their NMF procedure will be a practical tool for structure-finding that will probably have a significant effect on future work in the area. ■

Bartlett W. Mel is in the Department of Biomedical Engineering, University of Southern California, MC 1451, Los Angeles, California 90089, USA.
e-mail: mel@inc.usc.edu

1. Lee, D. D. & Seung, H. S. *Nature* **401**, 788–791 (1999).
2. Bishop, C. M. *Neural Networks for Pattern Recognition* (Oxford Univ. Press, 1995).
3. Hinton, G. E. & Sejnowski, T. J. *Unsupervised Learning: Foundations of Neural Computation* (MIT Press, Cambridge, Massachusetts, 1999).
4. Penev, P. S. & Atick, J. J. *Comp. Neural Syst.* **7**, 477–500 (1996).
5. Biederman, I. *Psychol. Rev.* **94**, 115–147 (1987).

erratum Due to an editorial oversight, the picture accompanying the article "Neurobiology: Cognition by a mini brain" (*Nature* **400**, 718–719; 1999), did not carry the correct attribution. It should have been credited to Karlheinz Rein of the Lehrstuhl für Genetik, Biozentrum, Würzburg.

Daedalus

Setting and upsetting

Last week Daedalus presented his 'Wet Cement', a hygroscopic concrete which set but never dried out. A building made from it could creep rather than crack while settling into the soil. He is now taking the idea to extremes. His new, even more advanced cement does not merely creep under load; impact liquefies it completely. The result is the first thixotropic, re-usable concrete.

Anyone who has wiggled a foot in damp beach sand, stirring it into a mobile porridge, will understand the principle. A pile of sand can resist steady compression; but no such pile is fully compact. A few grains take all the load, leaving the rest essentially free and unloaded. Under sudden shock or stirring, the grains shift into a denser packing which occupies a smaller volume. The water around them, however, cannot shrink. The grains find themselves suspended in excess water under strong agitation. They slump into a strengthless fluidized mass.

So DREADCO's Thixocrete contains sand to carry the load, hygroscopic compounds and water-absorbing polymers to keep it wet, and a highly impure, mixed calciferous setting cement. As in ordinary cement, its crystals precipitate from the water to hold the sand grains in place. But the crystals, loaded with impurities and dislocations, are extremely brittle. They shatter to fragments under sudden shock or vibration, freeing the sand grains to fluidize in the water. Now small crystals preferentially dissolve in water whereas bigger ones crystallize from it (this is the 'ripening' action of a chemical precipitate). So when agitation ceases, new crystals slowly precipitate, and knit into a restraining matrix round the sand grains. The Thixocrete slowly sets again.

Thixocrete will transform the building trade. The new product will be made and transported not in slowly rotating mixers but in vigorously stirred tanks. Pumped into moulds or extruded along lines of bricks, it will set in minutes. Any mistake or design error will be easily rectified by a powerful vibrator or wrecker's ball, shocking the structure back to re-usable fluid Thixocrete. By the same token, a Thixocrete structure will be easily demolished. And an earthquake or bomb attack will not shatter it disastrously. Instead it will flow and slump. Falling pieces will be soft and plastic; escapers will wade to safety through them. And the deformed building will not be a hazardous, unstable ruin. It will slowly regain its strength.

David Jones

Hox genes and the making of sphincters

Hoxd genes are needed for the formation of sphincters to subdivide the developing gut.

As well as their shared role in the anteroposterior patterning of vertebrates¹, particular Hox gene complexes have evolved specific functions. For example, four consecutive genes of the *HoxD* cluster contribute to making proper digits^{2,3}. Functional studies of such traits are difficult because there is a high genetic redundancy⁴, which means that several gene inactivations *in cis* are needed to analyse these global regulations. We have used a *HoxD* mini-complex in mice to show that an overlapping, yet different, set of *Hoxd* genes contributes to the formation of a major anatomical and physiological subdivision of the gut, the ileocaecal sphincter, which divides the small intestine from the large bowel.

The mouse *HoxD* complex contains nine genes (Fig. 1a), in a genomic segment roughly 100 kilobase pairs long, that are important for the lower lumbosacral area, as well as for proper limb and urogenital development. In addition, five genes (*Hoxd4*, *Hoxd8*, *Hoxd9*, *Hoxd10* and *Hoxd11*) showed similar expression in the intestinal hernia^{5–7} (the part of the gut that develops early outside the abdominal cavity) (Fig. 1b–d). However, individual loss of functions failed to reveal a corresponding alteration.

We engineered a mini-complex in which all the *Hoxd* genes were deleted except for *Hoxd1* and *Hoxd3*, although *Hoxd3* was functionally impaired after its regulatory regions were removed. We determined the residual regulatory potential of this locus by introducing a *Hoxd11/lacZ* transgene, the regulation of which was known when it was integrated either randomly⁸ or within the complex itself^{3,9}. From day 11 onwards, expression of *lacZ* in the mini-complex was strong in hernial gut as well as in limbs (Fig. 1d). Expression was also detected more anteriorly than for *Hoxd3* (ref. 8), as it reached the midbrain–hindbrain boundary (Fig. 1d, arrow). Unlike other *Hoxd* genes, it was widely expressed in derivatives of the first branchial arch (Fig. 1c), as well as in cells derived from the second rhombomere.

Gut dissection revealed that the staining in the hernia matched the ileocaecal transition (Fig. 1e,f), which is the passage between the posterior ileum and the colon, defining the position of the caecum. Staining was strong in the mesoderm but absent from epithelial cells. In contrast, proximal to the sphincter, staining was found in epithelium but not in muscle layers (Fig. 1i,j). The transgene was also expressed in caudal gut, reflecting the function of posterior *Hoxd*

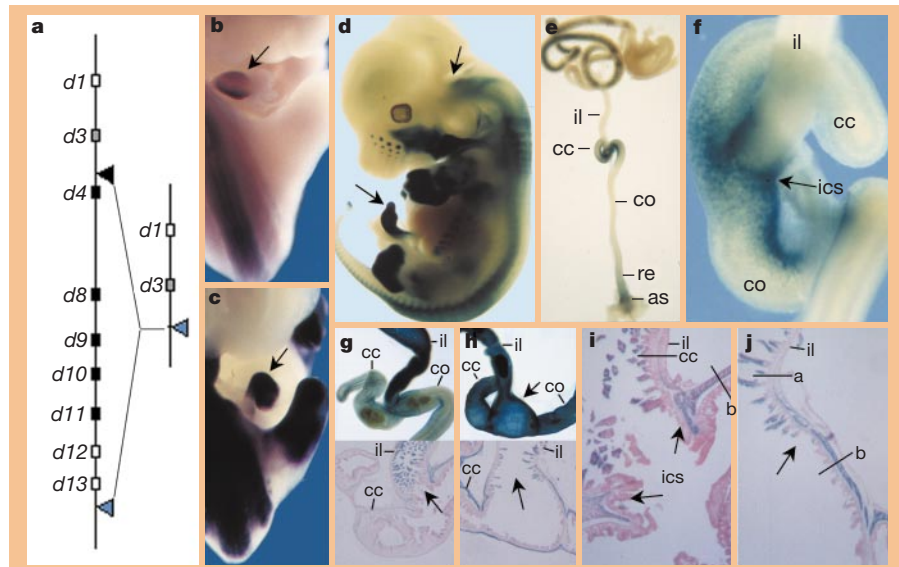


Figure 1 *Hoxd* genes and subdivision of the gut. **a**, Genes expressed in the intestinal hernia are shown in black. Arrowheads, loxP sites introduced by homologous recombination. Blue arrowheads, the *Hoxd11/lacZ* transgene. Treatment with Cre recombinase produced a mini-complex containing only *Hoxd1*, *Hoxd3* and the reporter transgene. **b**, **c**, Expression of *Hoxd4* (**b**) and *Hoxd11* (**c**) in the hernia of fetuses at embryonic day 11 (E11). **d**, Transgene expression at E13 reached the hindbrain–midbrain boundary and was strong in the hernia (arrow). **e**, Dissected gut with staining visible around both posterior sphincters. **f**, Enlargement of the ileocaecal region with strong staining around the sphincter, whereas the ileum, caecum and anterior colon were not stained. **g–j**, Gastrointestinal malformation in deleted homozygous mice (**h**). Control mice had a well formed ileocaecal transition (**g**, **i**), but the valve was absent in a homozygous mouse and the lower ileum, caecum and colon were distended (**h**, **j**). Corresponding histological analyses at six days of age are shown with X-gal staining and eosin counterstain, with enlargements of the ileocaecal areas (**i**, **j**). In control mice, a transition was observed in *lacZ* expression from the muscle layers to the epithelium (lines a and b) at the level of the sphincter. In homozygous mice, the layers were disorganized and no thickening was observed even though the transition was maintained (labelled a and b). Abbreviations: il, ileum; cc, caecum; co, colon; re, rectum; as, anal sphincter; ics, ileocaecal sphincter.

genes in the formation of the ano-rectal sphincter¹⁰. After the first week, mice homozygous for the mini-complex had retarded development, a third of them reaching less than half their normal body weight, and most of these died within two weeks.

The ileocaecal valve (the junction between the large and the small bowel) is normally marked by a thickening of the circular muscle coat, the ileocaecal sphincter (Fig. 1g,i). All homozygous mice with the *Hoxd* deletions lacked this sphincter, having instead a continuous transition from the lower ileum to the colon (Fig. 1h,j). At the ileocaecal transition, the smooth-muscle layer was thin and disorganized in homozygotes, leading to the absence of a sphincter (Fig. 1i,j). Analysis of the upper gut revealed signs of aberrant cell differentiation in the pyloric region of the stomach where ectopic islands of alkaline phosphatase-positive cells were found in the epithelium. Alterations to smooth muscle, as well as epithelial specification, may therefore cause the gastrointestinal defects that result in high postnatal lethality.

These results, along with alterations to the anal sphincter in *Hoxd13* and *Hoxd12*

mutants¹⁰, indicate that *Hoxd* genes are required to set up physiological constrictions along the previously unsubdivided gut mesoderm^{10–12}. In the absence of *Hoxd* function, mice lack sphincters. Mutations in homeotic genes in *Drosophila* can modify the pattern of muscle constrictions in midgut mesoderm¹³, indicating an ancestral role for these genes in assigning specificities along a primitive digestive system.

The requirement for several neighbouring *Hoxd* genes to make intestinal sphincters indicates that, during evolution, the existence of gene clusters probably favoured the functional recruitment of several genes at once, making this process the rule rather than the exception. Thus, although enhancer sharing may be a strong constraint for the maintenance of Hox genes together, the possibility that this regulatory parsimony is a consequence rather than a cause of clustering should not be overlooked.

József Zákány, Denis Duboule

Department of Zoology and Animal Biology,
University of Geneva, Sciences III,
Quai Ernest Ansermet 30,
1211 Geneva 4, Switzerland
e-mail: denis.duboule@zoo.unige.ch

- Krumlauf, R. *Cell* **78**, 191–201 (1994).
- Dollé, P. *et al. Nature* **342**, 767–772 (1989).
- van der Hoeven, F. *et al. Cell* **85**, 1025–1035 (1996).
- Rijli, F. & Chambon, P. *Curr. Opin. Gen. Dev.* **7**, 481–487 (1997).
- Condie, B. G. & Capecchi, M. R. *Development* **119**, 579–595 (1993).
- Izpisua-Belmonte, J. C. *et al. Development* **110**, 733–746 (1991).
- Dollé, P. & Duboule, D. *EMBO J.* **8**, 1507–1515 (1989).
- Gérard, M., Duboule, D. & Zákány, J. *EMBO J.* **12**, 3539–3550 (1993).
- Zákány, J. & Duboule, D. *Nature* **384**, 69–71 (1996).
- Kondo, T. *et al. Development* **122**, 2651–2659 (1996).
- Roberts, D. J. *et al. Development* **121**, 3163–3174 (1995).
- Sekimoto, T. *et al. Genes Cells* **3**, 51–64 (1998).
- Bien, M. *Trends Genet.* **10**, 22–26 (1994).

Evolution

Dicyemids are higher animals

Dicyemids, which are microscopic parasites of squids and octopuses, have among the simplest body plans of all multicellular animals. They lack body cavities and almost all the organs that characterize animals, such as a gut or nervous system, and their development proceeds without germ layers and gastrulation. The adult body consists of a solitary axial cell surrounded by a single layer of 10–40 ciliated outer cells. Here we use information from Hox gene sequences to investigate the phylogenetic affinities of dicyemids, and conclude that dicyemids are lophotrochozoans that have secondarily lost many morphological characters, so the simplicity of their body plan is not a primitive condition.

Because of their simple body plan, dicyemids have long been the subject of phylogenetic controversy. They comprise most of the Mesozoa, a name that indicates a level of complexity intermediate between protists and Metazoa¹. Alternatively, they have been suggested to be derived from parasitic platyhelminths² (dicyemids are obligate symbionts found in the renal sacs of cephalopod molluscs) or placed within the protists³. Data from 18S-ribosomal DNA indicate that dicyemids may be related to triploblasts (bilaterians), although these analyses do not resolve whether dicyemids are primitive or degenerate triploblasts⁴. Dicyemids are recognized as primitive animals, although their phylogenetic placement relative to other basal taxa such as Porifera, Placozoa and Cnidaria is unclear.

We isolated genomic and complementary DNA clones for a Hox gene, *DoxC*, from the dicyemid mesozoan *Dicyema orientale*. We used high-stringency Southern-blot analysis and the polymerase chain reaction (PCR) with *DoxC*-specific primers to confirm that *DoxC* is derived from the dicyemid genome: both methods detected the gene in dicyemid DNA but not in host squid DNA (data not shown). The homeo-domain sequence indicates that *DoxC* is a member of the 'middle' group⁵ of Hox (or Hox-like) genes, and identity is highest with *Antp* and its orthologues. The middle group of Hox genes has only been reported from triploblasts; no cnidarian genes fall into this group⁶.

A diagnostic peptide motif is encoded immediately carboxy-terminal to the homeo-domain (Fig. 1). This 'spiral peptide'⁷ (also called the Lox5 peptide⁸) has only been reported from the *Antp* orthologue in lophotrochozoans⁹, a group that includes annelids, ribbon worms, brachiopods and planarians (recently assigned to this clade). The peptide motif is not present in any other Hox proteins, including the *Antp* orthologue of ecdysozoan protostomes (which include nematodes, arthropods and priapulids), the Ubx/abd-A-type proteins of ecdysozoans or lophotrochozoans, or any proteins from deuterostomes or cnidarians. The presence of the spiral peptide indicates that *DoxC* is the dicyemid orthologue of *Lox5* from leeches, polychaetes and brachiopods, *LsHox6* from ribbon worms and *Dthox-C*, *Dthox-E* and *Pnox-7* from planaria. It also implies that dicyemid mesozoa are not basal and primitive animals and should not be excluded from Metazoa.

Rather, our data argue that dicyemids

are members of the Lophotrochozoa and are related to phyla such as platyhelminths, molluscs, nemerteans, brachiopods and annelids. We conclude that dicyemids are secondarily simplified from higher protostome animals and that their body plan is enormously reduced as a result of their endoparasitic lifestyle. Along with the enigmatic Myxozoa, they represent one of the most extreme cases of secondary reduction of body-plan complexity.

Mari Kobayashi*†, Hidetaka Furuya‡, Peter W. H. Holland*

*School of Animal and Microbial Sciences, University of Reading,

Whiteknights, Reading RG6 6AJ, UK
e-mail: p.w.h.holland@reading.ac.uk

†Department of Zoology, Graduate School of Science, Kyoto University, Kyoto 606, Japan

‡Department of Biology, Graduate School of Science, Osaka University, Toyonaka 560, Japan

- van Beneden, E. *Bull. Acad. R. Belg. Classe Sci.* **42**, 35–97 (1876).
- Stunkard, H. Q. *Rev. Biol.* **29**, 230–244 (1954).
- Cavalier-Smith, T. *Microbiol. Rev.* **57**, 953–994 (1993).
- Katayama, T., Wada, H., Furuya, H., Satoh, N. & Yamamoto, M. *Biol. Bull.* **189**, 81–90 (1995).
- Brooke, N.M., Garcia-Fernández, J. & Holland, P. W. H. *Nature* **392**, 920–922 (1998).
- Martinez, D. E., Bridge, D., Masuda-Nakagawa, L. M. & Cartwright, P. *Nature* **393**, 748–749 (1998).
- Bayasas, J. R., Castillo, E. & Salo, E. *Dev. Genes Evol.* **208**, 467–473 (1998).
- de Rosa, R. *et al. Nature* **399**, 772–776 (1999).
- Aguinaldo, A. M. *et al. Nature* **387**, 489–493 (1997).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

Insect behaviour

Male beetles attracted by females mounting

Intrasexual mounting is performed by males and females of many taxa¹, and female–female mounting occurs in insects, lizards, birds and mammals^{1,2}. Although the adoption by females of other male-like characters, such as mimicry of male colour patterns^{3–5}, is known to be advantageous, the benefits of female–female mounting have remained mysterious. Here we describe a pattern of female–female mounting in the beetle *Diaprepes abbreviatus* (Curculionidae) and demonstrate that it conveys a possible evolutionary advantage by providing a greater opportunity for the females to mate with larger males. This explanation may also apply to female intrasexual mounting in several other insect species.

We collected adult weevils from their food plants in Homestead, Florida, and observed them mounting females (Fig. 1). Female and male mounting was similar, and two females could easily be mistaken for a

spiral peptide consensus N..KLTGPG			
DoxC	HKRTQYTYRSQTLEKEF	HYNKYLTKRRRIEIAHLC	TERQIKIWPQNRMRKWKKEA
Dthox-C	NKRTQYTYTRHQTLEKEF	HFNKYLTKRRRIEIAHLC	TERQIKIWPQNRMRKWKKEA
Dthox-E	HKRSQYTYRYQLEKEF	HFNKYLTKRRRIEIAHLC	TERQIKIWPQNRMRKWKKEA
CTs-Lox5	-----	HYNKYLTKRRRIEIAHLC	TERQIKIWPQNRMRKWKKEA
Hr-Lox5	QKRTQYTYRYQLEKEF	YSNKYLTKRRRIEIAHLC	TERQIKIWPQNRMRKWKKEA
LsHox6	QKRTQYTYRYQLEKEF	HFNKYLTKRRRIEIAHLC	TERQIKIWPQNRMRKWKKEA
mab5	SKRTQYTYRSQTLEKEF	HYNKYLTKRRRIEIAHLC	TERQIKIWPQNRMRKWKKEA
Antp	RKRGRQYTYRYQLEKEF	HFNKYLTKRRRIEIAHLC	TERQIKIWPQNRMRKWKKEA
TgHox1	RKRGRQYTYRYQLEKEF	HFNKYLTKRRRIEIAHLC	TERQIKIWPQNRMRKWKKEA
AmphiHox6	RKRGRQYTYRYQLEKEF	HFNKYLTKRRRIEIAHLC	TERQIKIWPQNRMRKWKKEA
mouseHox6	GRGRQYTYRYQLEKEF	HYNKYLTKRRRIEIAHLC	TERQIKIWPQNRMRKWKKEA
mouseHox7	RKRGRQYTYRYQLEKEF	HFNKYLTKRRRIEIAHLC	TERQIKIWPQNRMRKWKKEA
mouseHox8	RRGRQYTYRYQLEKEF	LFNPKYLTKRRRIEIAHLC	TERQIKIWPQNRMRKWKKEA
		*****	*****

Figure 1 Comparison of deduced amino-acid sequences of the homeodomain (underlined) and carboxy-terminal flanking region for Hox proteins related to *Antp*. The dicyemid Hox gene, *DoxC*, encodes a spiral peptide motif assigning it to the Lophotrochozoa. Sequences shown for comparison derive from a platyhelminth (*Dthox-C* and *Dthox-E*), polychaete (*CTs-Lox5*), leech (*Hr-Lox5*), ribbon worm (*LsHox6*), nematode (*mab5*), *Antp* (fruitfly), sea urchin (*TgHox1*), cephalochordate (*AmphiHox6*) and mouse. The *DoxC* sequence has been deposited in the GenBank database, accession number AB030175. See Supplementary Information for the sequences studied and their accession numbers.



Figure 1 Female-female mounting in *Diaprepes abbreviatus*. The mounting female positions herself along the mounted female's back and extends her ovipositor as she does during oviposition, so it touches the lower posterior part of the mounted female.

pair in copula. Females showed no obvious response to mounting and remained together for up to 17.0 min (mean, 9.6 min) in the laboratory, compared with more than 10 hours when males mounted females⁶. Pairing ended either when the female dismounted (17 of 44 cases) or when a male contacted a female (27 of 44), at which point he was equally likely to mate with either female (13 males mated the mounting female, 11 mated the mounted female, and three males mated neither).

D. abbreviatus males and females are monomorphic, but males are slightly smaller than females (50% overlap in elytron length: male 8.7 ± 1.1 mm, female 10.3 ± 1.2 mm; *t*-test, d.f. = 812, $P < 0.001$). Males are attracted by volatile chemicals and pheromones but have difficulty discriminating aggregating females⁶, and often mount other males or a copulating male-female pair. In the absence of reliable cues, males search either for larger individuals, as these are more likely to be females, or for mating couples, as one individual is likely to be female. When a male approaches a mating couple, he pushes the mating male off the female's back, and larger males are more successful than smaller ones at such takeovers⁷.

We considered that mounting females might be mistaken for large mating males and so attract only large males that can win in male-male competition. We therefore tested the behaviour of large and small males by presenting them individually with five small and five large dead females glued in the mounting position on the backs of ten live, medium-sized females. As predicted, large males were attracted more to large pairs than to small pairs ($65.5 \pm 10.7\%$ and $34.5 \pm 10.7\%$, respectively; $G = 29.7$, d.f. = 8, $n = 80$, $P < 0.05$), and small males were more likely to approach small pairs than large ones ($71.7 \pm 8.2\%$ and $28.3 \pm 8.2\%$, respectively; $G = 16.07$, d.f. = 8, $n = 80$, $P < 0.05$). By mating preferentially with larger males, females may benefit either through 'good genes' or directly by obtaining resources, as materials are transferred from males to females during mating⁷.

Another possible explanation for female-female mounting is that the females are mimicking mating males to reduce male sexual harassment^{4,8,9}. However, this hypothesis was rejected because males were more attracted to 'mating' (glued) females ($70.8 \pm 10.9\%$) than to individual females ($29.2 \pm 10.9\%$; $G = 23.18$, d.f. = 8, $P < 0.05$). Alternatively, females may mount one another in order to appear larger¹⁰ and thus obtain better mates, as males prefer to mate with larger, more fecund females. However, small males are not attracted by mounting females, so this idea can also be rejected.

Our explanation is consistent with all known cases of naturally occurring female-female mounting in insects¹, as males have difficulty distinguishing females and so seek females in copulating pairs. By mounting one another, females can increase their opportunities to mate with large males.

Ally R. Harari*, H. Jane Brockmann

Department of Zoology, University of Florida, Gainesville, Florida 32611-8525, USA

*Present address: The Volcani Center, Department of Entomology, Beit Dagan 50250, Israel
e-mail: aharari@netvision.net.il

1. Bagemihl, B. *Biological Exuberance: Animal Homosexuality and Natural Diversity* (St Martin's Press, New York, 1999).
2. Fang, J. & Clements, L. G. *Anim. Behav.* **57**, 545-555 (1999).
3. Vane-Wright, R. I. in *The Biology of Butterflies* (eds Vane-Wright, R. I. & Ackery, P. R.) 251-253 (Princeton Univ. Press, New Jersey, 1989).
4. Robertson, H. M. *Anim. Behav.* **33**, 805-809 (1985).
5. Clarke, C., Clarke, F. M. M., Collins, S. C., Gill, A. C. L. & Turner, J. R. G. *Syst. Entomol.* **10**, 257-283 (1985).
6. Harari, A. R. & Landolt, P. J. *J. Chem. Ecol.* **23**, 857-868 (1997).
7. Harari, A. R., Handler, A. M. & Landolt, P. J. *Anim. Behav.* (in the press).
8. Scott, D. *Proc. Natl Acad. Sci. USA* **83**, 8429-8433 (1986).
9. Cordero, A., Santolamazza Carbone, S. & Utzeri, C. *Anim. Behav.* **55**, 185-197 (1998).
10. Beehler, B. M. & Foster, M. S. *Am. Nat.* **131**, 203-219 (1988).

Environmental toxins

Exposure to bisphenol A advances puberty

Plastics and pesticides are examples of products that contain oestrogenic endocrine-disrupting chemicals, or EEDCs, which can interfere with mammalian development by mimicking the action of the sex hormone oestradiol¹. For instance, the exposure of developing rodents to high doses of EEDCs advances puberty and alters their reproductive function². Low environmental doses of EEDCs may also affect development in humans³. Effects have become apparent in humans over the past half century that are consistent with those seen in animals after exposure to high doses of EEDCs, such as an increase in genital abnormality in boys⁴ and earlier sexual maturation in girls⁵. Here we show that exposing female mouse fetuses to an EEDC

at a dose that is within the range typical of the environmental exposure of humans alters the postnatal growth rate and brings on early puberty in these mice.

Oestrogen is a hormone that interacts with other steroids to regulate the normal development of the reproductive system and other tissues. Bisphenol A, a compound that was initially synthesized as a chemical oestrogen⁶, is now used as the monomer for the production of polycarbonate plastic products such as baby bottles. Bisphenol A leaches out of such products at a rate that increases with repeated use⁷.

Pregnant CF-1 mice ($n = 21$ per treatment group) were fed either oil (vehicle) or bisphenol A dissolved in oil at a dose equivalent to that typically found in the environment ($2.4 \mu\text{g}$ per kg), on days 11 to 17 of gestation⁸. Pups were delivered by caesarean section on day 19 to determine their position in the uterus and were reared by untreated foster mothers. The intrauterine position determines fetal hormone levels because endogenous sex steroids are transported from one fetus to another⁹. Mouse fetuses positioned between two males (Fig. 1, 2M) are exposed to the lowest levels of oestradiol, fetuses located next to female fetuses (0M) are exposed to the highest, and females next to one male (1M) are exposed to an intermediate amount¹⁰.

At weaning on postnatal day 22, females treated with bisphenol A were significantly heavier than control females (Fig. 1a), although they had a similar body weight at birth: relative to controls from the same intrauterine position, the weight of 0M females was increased by 22% and that of 1M females was increased by 9%; 2M females were unaffected (Fig. 1d). The findings were virtually identical for male siblings.

On postnatal day 26, females were housed individually but near to males to provide a submaximal level of pheromonal stimulation¹¹. After vaginal opening, daily vaginal smears were examined for the presence of completely cornified epithelial cells (a sign of first vaginal oestrus). We found that prenatal treatment with bisphenol A significantly reduced the number of days between vaginal opening and first vaginal oestrus, which is highly correlated with first postpubertal ovulation¹² (Fig. 1c), in 0M females but not in 2M females (Fig. 1f), based on analysis of covariance adjusted for body weight at weaning.

Prenatal exposure to a dose of bisphenol A comparable to levels found in the environment therefore altered postnatal growth rate and reproductive function in female mice, although individual differences in endogenous oestradiol resulting from natural variation influenced the responsiveness of the females to bisphenol A.

There is significant variability in human and animal populations in responsiveness

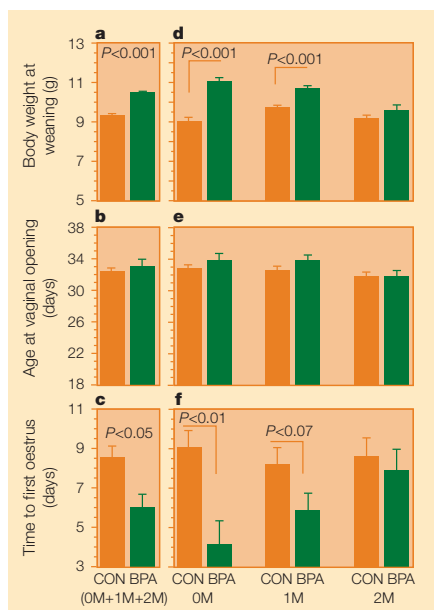


Figure 1 Mean ($\pm$ s.e.m.) body weight at weaning, age at vaginal opening, and interval between vaginal opening and first vaginal oestrus. Data are for all females combined (a–c) and as a function of intrauterine position (d–f). Body weight at weaning includes data for all females surviving to weaning from control (CON, orange bars) and bisphenol A-exposed (BPA, green bars) litters. All data were adjusted for litter membership to control for maternal effects. Vaginal opening and interval data were also corrected by analysis of covariance for body weight at weaning. 2M, located between two males fetuses; 1M, located next to one male fetus; OM, located next to female fetuses. Wean weight was calculated on 41 OM, 47 1M and 23 2M control females, and 20 OM, 43 1M and 12 2M bisphenol A-treated females. Vaginal opening and interval data were calculated on 19 OM, 20 1M and 19 2M control females, and on 19 OM, 21 1M and 11 2M bisphenol A-treated females. We attempted to include females from each intrauterine position from each litter, but some litters did not contain a 2M female.

to hormones and EEDCs. Our findings indicate that one source of this variability may be the amount of endogenous sex hormones. These vary among individual human fetuses and are influenced by a variety of factors, including whether it is a first pregnancy, the size of the placenta, and whether there is just one fetus or twins^{13,14}.

Very small increases in the level of endogenous oestradiol may substantially increase the sensitivity of fetuses to EEDCs consumed by pregnant women, so some fetuses may be at particularly high risk for a wide array of abnormalities and diseases. Our findings emphasize the need for studies to examine the relationship between maternal exposure to endocrine disrupters and subsequent health effects in the offspring.

Kembra L. Howdeshell*, Andrew K. Hotchkiss†, Kristina A. Thayer*, John G. Vandenbergh†, Frederick S. vom Saal*

*Division of Biological Sciences, University of Missouri, Columbia, Missouri 65211, USA
e-mail: vomsaalf@missouri.edu

†Department of Zoology,

North Carolina State University,
Raleigh, North Carolina 27695-7617, USA

- Colborn, T., vom Saal, F. S. & Soto, A. M. *Environ. Health Perspect.* **101**, 378–384 (1993).
- Gray, L. E. Jr. in *Chemically Induced Alterations in Sexual and Functional Development* Vol. 21, *The Wildlife/Human Connection* (eds Colborn, T. & Clement, C.) 203–230 (Princeton Scientific, New Jersey, 1992).
- vom Saal, F. S. & Sheehan, D. M. *Forum* **13**, 11–18 (1998).
- Paulozzi, L. J., Erickson, J. D. & Jackson, R. J. *Pediatrics* **100**, 831–834 (1997).
- Herman-Giddens, M. E. *et al. Pediatrics* **99**, 505–512 (1997).
- Dodds, E. C. & Lawson, W. *Nature* **137**, 996 (1936).
- Takao, Y. *et al. J. Health Sci.* **45**, 39 (1999).
- Nagel, S. C. *et al. Environ. Health Perspect.* **105**, 70–76 (1997).
- Even, M. D., Dhar, M. G. & vom Saal, F. S. *J. Reprod. Fertil.* **96**, 709–716 (1992).
- vom Saal, F. S. *J. Anim. Sci.* **67**, 1824–1840 (1989).
- Vandenbergh, J. G. *Endocrinology* **84**, 658–660 (1969).
- vom Saal, F. S., Pryor, S. & Bronson, F. H. *J. Reprod. Fertil.* **62**, 33–37 (1981).
- Panagiotopoulou, K. *et al. Cancer Causes Control* **1**, 119–124 (1990).
- Hsieh, C.-C. *et al. Am. J. Epidemiol.* **136**, 1321–1326 (1992).

Global warming

Solar variability and the Earth's climate

Lockwood *et al.*¹ recently presented some intriguing new evidence of solar variability, but Parker's accompanying News and Views article² gave an exaggerated and misleading picture of the potential effects on terrestrial climate. This picture is at variance with both the evidence³ and a public statement by Lockwood himself, reported in ref. 4.

As Parker mentions, the only available direct measurements show a variation of just 0.15% in solar irradiance *S* over one solar magnetic cycle of 11 years. Greater variations on longer timescales are possible, but other reports^{5,6} give figures that are less than the 0.5% that Parker quotes. Even $\Delta S/S \approx 0.5\%$ implies global mean radiative forcing: $(\Delta S/4)(1 - \alpha) \approx 1 \text{ W m}^{-2}$, where $\alpha \approx 0.3$ is Earth's albedo. For comparison, the Intergovernmental Panel on Climate Change (IPCC) estimates the change in global mean radiative forcing from 1861 to 1990 as 2.0 to 2.8 W m^{-2} from greenhouse gases, -0.2 to -2.3 W m^{-2} from aerosols, and 0.1 to 0.5 W m^{-2} from solar variability^{7,8}. There has been much speculation about additional solar effects but, with regard to Parker's remark about cosmic-ray effects on clouds, we note that the cited work⁹ claims that this effect would only amplify solar radiative forcing to less than 2 W m^{-2} .

We are surprised by Parker's suggestion that solar brightening is responsible not only for the observed increase in twentieth-century surface temperatures, but also for increased carbon dioxide in "the same way that a carbonated drink expels most of its CO_2 if warm". Given the observed surface temperature increase of about 1 K, a simple calculation¹⁰ reveals that this effect would increase equilibrium CO_2 by less than 5%,

which is much less than the 30% increase recorded since the beginning of the industrial revolution. Records of ^{13}C and ^{14}C isotopes relative to ^{12}C also support the idea that changes in atmospheric CO_2 are primarily due to the burning of fossil fuel⁸.

Although solar effects on this century's climate may not be negligible, quantitative considerations imply that they are small relative to the anthropogenic release of greenhouse gases, primarily carbon dioxide. Figure 8.4 of the IPCC 1995 assessment report³ (also shown in refs 7, 11) makes that point clearly, even though it assumes solar variability near the upper end of its uncertainty range⁷.

On longer timescales, it is not a "historical fact" that climate "responds to variations of the Sun's magnetic activity, with substantial warming and cooling with the rise and fall of activity over the centuries". Several recent palaeoclimate compilations^{12,13} indicate only that weak responses to putative solar variations have occurred over the past 500–1,000 years. The data suggest that global mean temperatures during the 1990s may have been the warmest of the millennium. Together with climate model simulations (which, contrary to Parker, quantify the combined effects of cloud, winds, ocean currents, solar effects and anthropogenic effects^{14–16}, however imperfectly), the data indicate that there is probably a strong anthropogenic component to twentieth-century global warming.

Martin I. Hoffert*, Ken Caldeira†, Curt Covey†, Philip B. Duffy†, Benjamin D. Santer†

*Department of Physics,
New York University, New York,
New York 10003, USA
†Atmospheric Science Division,
Lawrence Livermore National Laboratory,
Livermore, California 94550, USA

- Lockwood, M., Stamper, R. & Wild, M. N. *Nature* **399**, 437–439 (1999).
- Parker, E. N. *Nature* **399**, 416–417 (1999).
- Santer, B. D. *et al. in Climate Change 1995: The Science of Climate Change* (eds Houghton, J. T. *et al.*) 406–443 (Cambridge Univ. Press, 1996).
- Pearce, F. *New Scientist* p. 5 (5 June 1999).
- Lean, J., Skumanich, A. & White, O. *Geophys. Res. Lett.* **19**, 1591–1594 (1992).
- Hoyt, D. V. & Schatten, K. H. *J. Geophys. Res.* **98**, 18895–18906 (1993).
- Harvey, L. D. D. *et al. An Introduction to Simple Climate Models used in the IPCC Second Assessment Report*, IPCC Technical Paper II (WMO, Geneva, 1997).
- Schmied, D. *et al. in Climate Change 1995: The Science of Climate Change* (eds Houghton, J. T. *et al.*) 65–139 (Cambridge Univ. Press, 1996).
- Svensmark, H. & Friis-Christensen, E. *J. Atmos. Terrest. Phys.* **59**, 1225–1232 (1997).
- Walker, J. C. G. *Numerical Adventures with Geochemical Cycles* (Oxford Univ. Press, 1991).
- Covey, C. & Hoffert, M. I. *Clim. Change* **37**, 387–390 (1997).
- Mann, M. E. & Bradley, R. S. *Geophys. Res. Lett.* **26**, 759–762 (1999).
- Briffa, K. R. & Osborn, T. J. *Science* **284**, 926–927 (1999).
- Cubasch, U. *et al. Clim. Dynam.* **13**, 757–767 (1997).
- Wigley, T. M. L., Smith, R. L. & Santer, B. D. *Science* **282**, 1676–1679 (1998).
- Tett, S. F. B. *et al. Nature* **399**, 569–572 (1999).

***En passant* neurotrophic action of an intermediate axonal target in the developing mammalian CNS**

Hao Wang* & Marc Tessier-Lavigne

Howard Hughes Medical Institute, Departments of Anatomy and of Biochemistry and Biophysics, University of California, San Francisco, California 94143-0452, USA

During development, neurons extend axons to their targets, then become dependent for their survival on trophic substances secreted by their target cells. Competition for limiting amounts of these substances is thought to account for much of the extensive naturally-occurring cell death that is seen throughout the nervous system. Here we show that spinal commissural neurons, a group of long projection neurons in the central nervous system (CNS), are also dependent for their survival on trophic support from one of their intermediate targets, the floor plate of the spinal cord. This dependence occurs during a several-day-long period when their axons extend along the floor plate, following which they develop additional trophic requirements. A dependence of neurons on trophic support derived *en passant* from their intermediate axonal targets provides a mechanism for rapidly eliminating misprojecting neurons, which may help to prevent the formation of aberrant neuronal circuits during the development of the nervous system.

Classic studies by Hamburger, Levi-Montalcini and others showed that many neuronal populations in vertebrates are initially generated in excess numbers during embryonic development, and that supernumerary neurons are subsequently eliminated by naturally-occurring neuronal cell death (reviewed in refs 1, 2). The massive death of excess neurons usually occurs soon after the neurons' axons reach their targets, and appears to result from a dependence of the neurons on trophic substances secreted by their target cells, which are present in limiting amounts and for which the neurons compete^{1,2}. This competition is thought to provide a mechanism for matching the size of the presynaptic neuronal population to the size of the target population, and has also been suggested to provide a means for eliminating misprojecting neurons, which would not have access to the target-derived trophic support.

With the identification of many neurotrophic factors, evidence has started to accumulate that neurons that have reached their targets may receive trophic support not just from their target cells but also from other cellular sources, such as those located near their cell bodies or axons (reviewed in refs 1, 2). There is also evidence that some differentiated neurons may have trophic requirements before they reach their target fields (reviewed in ref. 2, and see Discussion below). Because many developing axons, particularly those of long projection neurons, navigate to their final targets in a stepwise fashion from one intermediate target to the next, we were interested in testing whether neurons derive trophic support from their intermediate targets as their axons navigate in the vicinity of those targets. A requirement for intermediate target-derived trophic support would provide a means of eliminating misprojecting neurons soon after their axons become misrouted, and could therefore contribute to the accuracy of neuronal projection patterns.

Trophic action of the floor plate

To study potential trophic effects of an intermediate axonal target, we focused on a well characterized group of long projection neurons, the commissural neurons in the dorsal spinal cord (DSC), which send axons to distant targets by first projecting to

an intermediate target, the floor plate at the ventral midline of the spinal cord (Fig. 1a). In the rat, commissural neurons are born between embryonic day 11 (E11) and E13, and their axons take about a day to reach the floor plate³. These axons cross the midline at the floor plate and then turn sharply rostrally and travel longitudinally along the floor plate for several days to different axial levels in the spinal cord and brain, where they leave the midline to seek out their final targets³. Commissural axons are directed to the ventral midline at least partly by the floor plate-derived chemo-attractant netrin-1, which can stimulate the outgrowth of commissural axons and orient their growth *in vitro*⁴⁻⁸. The floor plate does not, however, appear to influence the survival of commissural neurons during this initial period of growth to the ventral midline⁵.

Because commissural axons grow along the floor plate for an extended period after reaching it, we examined whether the floor plate affects commissural neuron survival during this later period. In initial experiments using cultures of dissociated E13 rat DSC, we found that virtually all plated cells (95%) survived for the first 24 h, but after a further 48 h all had died (100%); when grown with floor plate-conditioned medium (FPCM), however, over 40% survived (data not shown). These results are consistent with a trophic effect of the floor plate on commissural neurons, but the absence of markers to distinguish these neurons from other DSC cells at these late stages prevented a firm conclusion. To circumvent this problem, we used cultures of explanted DSC tissue, taking advantage of the ability to elicit the selective outgrowth of commissural axons from these explants. When explants of E13 rat DSC are cultured in collagen gels either with a piece of floor plate tissue, with FPCM or with netrin-1, axons extend from these explants into the collagen within 12 h, and by 24 h have projected extensively in thick bundles that are not observed with E13 DSC explants cultured alone⁷ (Fig. 1b, e, h); these axons are commissural axons, as assessed by specific marker expression^{7,9}. Beginning after around 40 h in culture, however, many bundles extending from explants cultured with netrin-1 alone began to exhibit membrane swellings and blebs, and by 48 h extensive blebbing of all bundles was observed (Fig. 1f, g). Blebbing of the few axons that extended from control explants was also seen at this time (Fig. 1c, d). In contrast, we saw no blebbing in explants cultured with FPCM (added after 24 h in culture) or cocultured with floor plate tissue (placed within

* Present address: Merck Research Labs, WP26-265, West Point, Pennsylvania 19486, USA

~300 μm of the explants), irrespective of whether the medium was supplemented with netrin-1 (Fig. 1i, j and data not shown; in subsequent experiments, all cultures were supplemented with netrin-1 at the start). Netrin-1 was not required for the effect, as medium conditioned by floor plate cells from netrin-1-deficient mice⁸ also prevented blebbing of the few axons that extended from control explants, and of the axon bundles elicited by netrin-1 added for just the first 24 h in culture (data not shown).

Blebs associated with axon bundles appeared to represent disintegrating axons rather than cells migrating from explants, which were largely confined to the proximal portion of axons (Fig. 2c, d). Consistent with this, blebbing was associated with a marked reduction in staining with an antibody against the axonal marker neurofilament (Fig. 3b, c, e, f and data not shown). Axon disintegration also correlated with extensive apoptotic death of cells within the explants, as assessed using the TdT-mediated dUTP nick end labelling (TUNEL) assay (Fig. 2e), whereas few TUNEL-positive nuclei were observed in cultures with FPCM (Fig. 2f). These findings indicate that commissural neurons in E13 DSC explants require trophic support starting after one day in culture, that floor

plate cells secrete a factor(s) that can provide this support, and that in the absence of this factor the cells rapidly die and their axons disintegrate. The DSC explants used here are also expected to give rise to other interneurons that do not extend axons along the floor plate^{3,5}. As these neurons are born a day or more later than commissural neurons³, and are not known to extend axons from explants in response to floor plate or netrin-1, our experiments do not address whether floor plate cells also affect their survival.

Age dependence of trophic requirement

To determine whether this trophic dependence is an intrinsic property of commissural neurons determined by developmental stage, we characterized the time course of axonal disintegration in explants from different ages. As commissural neurons are born from E11 to E13 (ref. 3) and extend axons from both E11 and E13 DSC explants in response to netrin-1 (ref. 7), we compared cell death in cultures of explants from these two ages. Swift and extensive disintegration of axon bundles occurred in both cases when explants were cultured alone or with netrin-1, and was suppressed in both cases by FPCM, but in the case of E11 explants disintegration started two days later than with E13 explants (Fig. 3a–c and data not shown). The simplest interpretation of this result is that commissural neurons acquire dependence on a floor-plate-derived trophic factor at a developmental stage corresponding to E14–15, and that they can develop this dependence either *in vivo* or *in vitro*. Consistent with developmental stage being a determinant of trophic dependence, we also found that FPCM could support survival of commissural neurons only for about three days after the onset of trophic dependence (to the equivalent of E17 in each case), regardless of the stage (E11, E12, E13 or E14) at which explants were cultured (Fig. 3d–f and data not shown).

Spatial distribution of trophic activity

Preliminary characterization showed that the active component in FPCM has a relative molecular mass greater than 10,000 ($M_r > 10K$) and is sensitive to trypsin treatment, indicating that

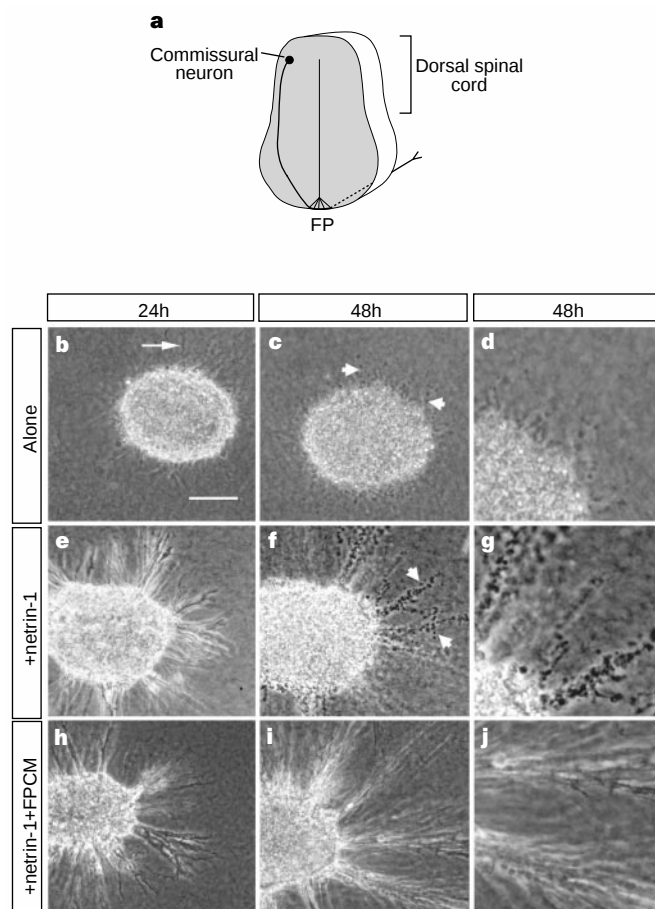


Figure 1 The floor plate (FP) secretes a diffusible factor(s) that suppresses the disintegration of commissural axons from E13 explants. **a**, Drawing of the E13 rat spinal cord showing the initial trajectory of commissural axons from cell bodies in the DSC to the FP at the ventral midline. After crossing the midline, commissural axons travel near the FP (dashed line) for several days. **b–j**, E13 rat DSC explants cultured in collagen gels after 24 h (**b**, **e**, **h**) or 48 h (**c**, **f**, **i**; higher magnification in **d**, **g**, **j**) ($n \geq 20$ explants for each condition). **b–d**, Few axons grew out by 24 h from explants cultured alone (arrows in **b**), and they disintegrated by 48 h (arrowheads in **c**). **e–g**, Explants cultured with 30 ng ml⁻¹ netrin-1 showed extensive outgrowth at 24 h (**e**), but all axon bundles disintegrated by 48 h (arrowheads in **f**). **h–j**, Axon bundles growing from explants cultured with 30 ng ml⁻¹ netrin-1 but with FPCM added at 24 h (**h**) showed no blebbing at 48 h. Scale bar: 100 μm in **b**, **c**, **e**, **f**, **h**, **i**; 50 μm in **d**, **g**, **j**.

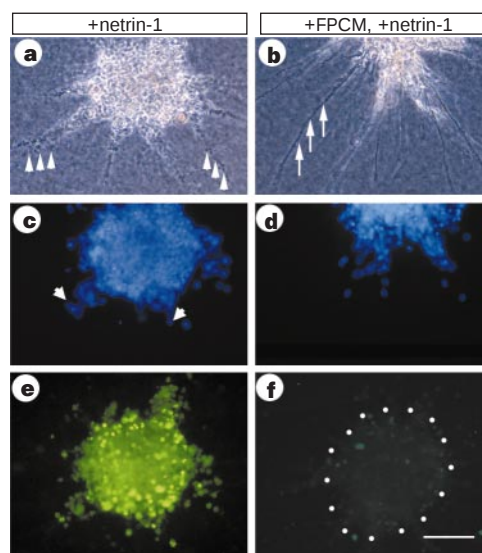


Figure 2 Axonal disintegration is associated with cell death within explants. **a**, **c**, **e**, E13 explant cultured with netrin-1 alone; **b**, **d**, **f**, a different explant cultured with both netrin-1 and FPCM. The extensive blebbing of axon bundles (arrowheads) observed in E13 explants cultured with netrin-1 for 48 h (**a**) is not associated with extensive migration of cells (visualized with DAPI (**c**, arrowheads)), but is associated with extensive apoptosis throughout the explant, as detected using the TUNEL assay (**e**). The suppression of axon blebbing by FPCM (**b**, arrows) appears to reflect a suppression of cell death (**f**; dots outline the explant) ($n = 10/10$ explants). Scale bar: 80 μm .

it is associated with a protein. We tested 27 defined factors (see Methods), including many known to possess neurotrophic activity, such as the neurotrophins nerve growth factor (NGF), brain-derived neurotrophic factor (BDNF) and neurotrophin-3 (NT-3). None of these factors possessed the activity either alone or in various combinations (see Methods). In the absence of an identification of the trophic factor(s), we determined its distribution in the spinal cord along the initial trajectory of commissural axons through bioassay of spinal cord fragments. Segments of E13 rat spinal cord tissue were subdivided into dorsal (D), intermediate (I), ventral (V) and floor plate (F) strips (Fig. 4, top). To quantify survival-promoting activity, fragments of these strips of roughly equal protein mass (Fig. 4 legend) were placed 300–350 μm from E13 DSC explants. After 44 h in culture, extensive disintegration was observed in 100% of control DSC explants, but not in any explants cultured with F fragments (0%). At this time, disintegration in cultures with V, I or D fragments was observed in ~40%, 70% and 80% of explants, respectively, with more extensive disintegration observed at 52 h, and virtually complete disintegration by 66 h (Fig. 4). In contrast, only a small amount of disintegration (in ~10% of explants) was observed at 52 and 66 h with F fragments. Thus, trophic activity is present all along the dorsoventral axis, but at much higher levels in the floor plate and in a decreasing ventral-to-dorsal gradient, reminiscent of the gradient of murine *netrin-1* transcripts⁸. The activity detected outside the floor plate may be

produced by these tissues or instead have diffused there from floor plate cells *in vivo* before explantation.

Local action of the floor plate at the axons

As the trophic activity is largely concentrated at the floor plate, for it to promote commissural neuron survival effectively *in vivo* we predict that it would have to act on commissural axons as they navigate the midline, with the trophic signal being retrogradely transported from the axons to their cell bodies. Retrograde transport of a trophic signal has been shown for NGF actions on sympathetic neurons using the 'Campanot chamber' system in which axons are made to grow through a grease-sealed partition to a separate chamber from that containing their cell bodies, so that axons can be exposed to NGF independently of their cell bodies¹⁰. The rate of commissural axon extension is, however, too slow to permit establishment of such partitioned cultures in the time frame during which the neurons can still be rescued by FPCM (data not shown). As an alternative approach, we took advantage of the ability of collagen gels to stabilize gradients of secreted factors, and sought to define experimental conditions in which axons and cell bodies are differentially exposed to limiting concentrations of trophic factor. For this, we first set up cultures of E13 DSC explants with small pieces of floor plate of defined size (Fig. 5) placed ~300 μm to one side, with netrin-1 added to elicit radial outgrowth. Efficient suppression of axon disintegration was observed at 52 h for only those bundles that grew within the quadrant closest to the floor plate tissue (Fig. 5a–d). As the cell bodies of axons that exit the explant are located within 20 μm of the exit point (as found in parallel DiI-labelling experiments; data not shown), in such cultures rescue of proximal quadrant axons could in principle be due to the factor diffusing to and acting at cell bodies in the proximal quadrant, or to the factor acting on the axons, or both. However, we obtained evidence that the factor must be acting on the axons in at least some of these cultures by focusing on those in which explants were exposed to a 'limiting concentration' of factor.

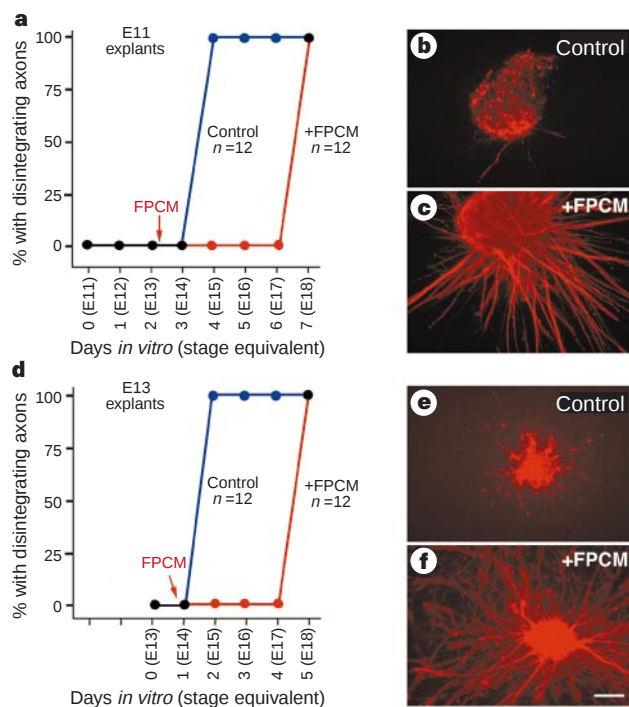


Figure 3 Developmental-stage dependence of axonal disintegration with or without FPCM. Axonal disintegration was a synchronous phenomenon, as extensive disintegration of all bundles in any given explant occurred within <24 h of the first signs of blebbing (data not shown). In these experiments FPCM was used at saturating concentration; under more limiting conditions (using small floor plate fragments) more asynchrony is observed (Figs 4 and 5). Cultures were performed with 30 ng ml⁻¹ (E13) or 300 ng ml⁻¹ (E11) netrin-1 to elicit axon outgrowth. **a, d**, Graphs show percent E11 (**a**) and E13 (**d**) explants with disintegrating axons at different stages (shown as days *in vitro* (DIV) and equivalent developmental stage) when cultured without FPCM (blue) or with FPCM (red) (where the curves overlap they are shown in black). **b, c**, Neurofilament staining of E11 explants fixed after 6 DIV shows few remaining axons in control cultures (**b**), in contrast to cultures with FPCM (**c**). **e, f**, Neurofilament staining of E13 explants cultured for 4 DIV without (**e**) or with (**f**) FPCM. Scale bar: 100 μm .

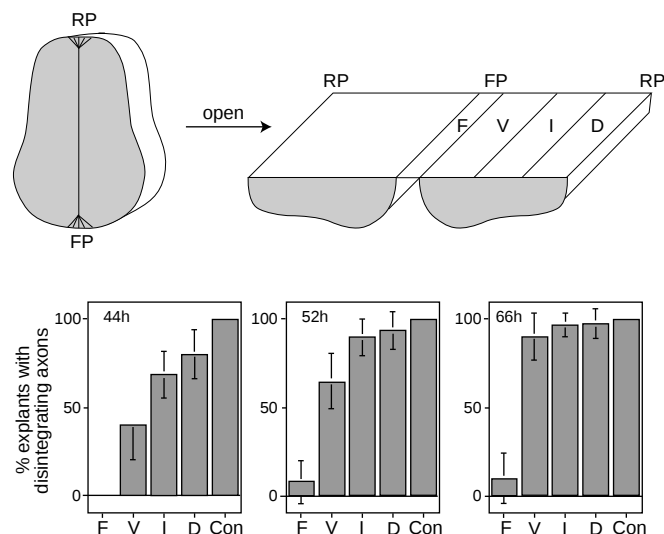


Figure 4 Evidence for a decreasing ventral-to-dorsal gradient of floor-plate-derived trophic activity in the E13 spinal cord. Top: Dissection of spinal cord segments (each segment was ~1/8 of the spinal cord). Segments were opened at the roof plate (RP) to generate an 'open book' preparation, and the spinal cord was divided into three equal strips (D (dorsal), I (intermediate) and V (ventral)) along the D–V axis. Strips of floor plate (F) were also isolated. Bottom: E13 DSC explants were obtained from a separate spinal cord, and positioned in a circle of radius ~300 μm around the following tissue fragments (of similar protein mass (see Methods)): 2 F strips, 1/3 of a V strip, or half of an I or a D strip. The percentage of DSC explants showing disintegration in the presence of these different tissues was quantified at 44 h, 52 h and 66 h ($n \geq 30$ explants per condition).

These explants are operationally defined as those in which all axon bundles that remained further than 300 μm from the floor plate disintegrated (19 explants met this criterion). In each of these explants, we then studied pairs of axon bundles that emanated from the same exit point from the explant (and whose cell bodies are therefore located close to one another), but in which one bundle stayed more than 170 μm from the floor plate, and the other grew to the vicinity of the floor plate. Twenty-one pairs of bundles from the 19 explants met this criterion (an example is shown in Fig. 5e). In all 21 cases the bundle that grew near the floor plate survived, whereas

the other one disintegrated. These results are not readily explained by differential effects of the trophic factor on the cell bodies that give rise to the two bundles, and instead indicate strongly that the factor acts on the axons to support survival under these limiting conditions. Further evidence for an action on axons came from examining cell bodies in limiting cultures (defined as above) using combined neurofilament and TUNEL staining. Neurofilament-positive axons that project to the vicinity of the floor plate were seen to originate from TUNEL-negative regions of the explants surrounded by TUNEL-positive cells in immediately adjacent regions ($n = 5$ explants; Fig. 5f–i and data not shown), consistent with the selective rescue of neurons whose axons project near the floor plate. These results again indicate that the concentration of trophic factor at the cell bodies is not sufficient to rescue neurons and that it must therefore be acting on the axons, at least under these limiting conditions.

Development of further trophic requirements

The observation that floor plate-derived trophic support is not adequate to support commissural neurons past a developmental stage equivalent to E17 (Fig. 3) indicates that these neurons might acquire a novel trophic dependence at the time that they leave the floor plate to find their final targets. Indeed, although a combination of neurotrophins (BDNF, NT-3 and NGF) had no effect at earlier stages (Fig. 6a, b), the disintegration of axons could be suppressed by these neurotrophins at a developmental stage equivalent to E16 (three days *in vitro* for E13 explants), and in the presence of these factors survival was extended by up to an additional two days, from E17 to E19 (Fig. 6c, d). FPCM was still required during this period, as adding neurotrophins without FPCM did not suppress axon disintegration (data not shown). Thus, by E17 commissural neurons appear to acquire an additional dependence on neurotrophins. Whether the neurons ever lose a requirement for the floor-plate-derived trophic activity has not been determined.

Discussion

Our results are consistent with a model in which commissural neurons acquire a dependence on trophic support from floor plate

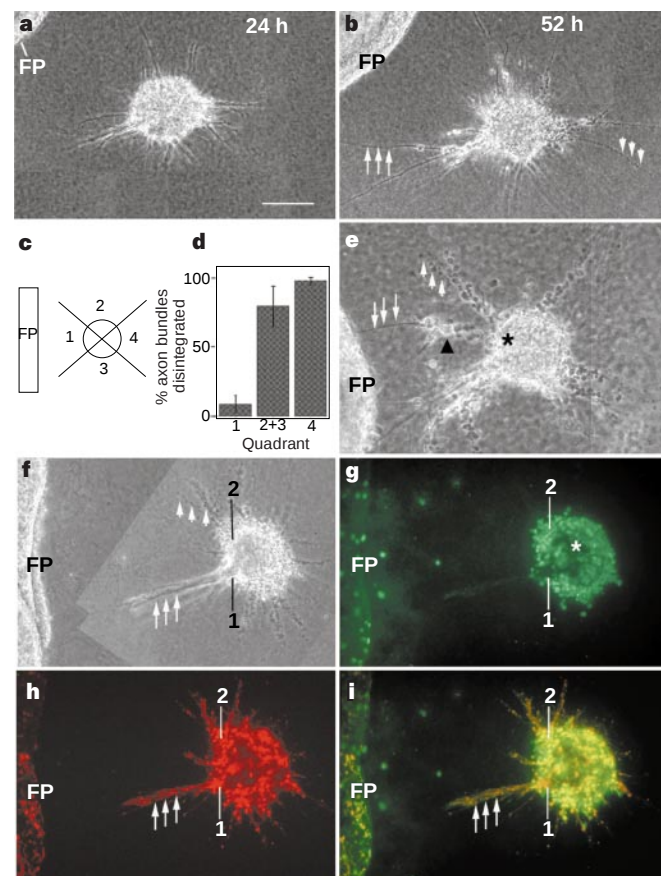


Figure 5 Local action of the floor plate-derived trophic activity on axons. **a–d**, Small pieces of floor plate (half an F strip; Fig. 4) placed $\sim 300 \mu\text{m}$ from E13 explants (grown with netrin-1 to elicit radial axon outgrowth, as seen at 24 h **a**) suppress axonal disintegration efficiently after 52 h only in the quadrant closest to the floor plate (**b**) (**d** shows quantification for quadrants 1–4 (defined in **c**); $n = 9$ explants). **e–i**, Explants exposed for 52 h to a ‘limiting concentration’ of trophic factor (see text). **e**, Pair of axon bundles (long and short arrows) emanating from the same spot (asterisk) in such an explant, in which one bundle approaches the floor plate (long arrows) and the other remains $>170 \mu\text{m}$ away (short arrows). Here, as in all other such bundle pairs (21 pairs in 19 explants), the distant bundle has disintegrated whereas the bundle closer to the floor plate is healthy, providing evidence for action of the trophic factor on the axons, not their cell bodies. (Note that rounded figures on the proximal portions of the axons (black triangle) are migrating cells, not axon blebs, as assessed in parallel experiments using DAPI (Fig. 2c, d).) **f–i**, Another explant exposed to a ‘limiting concentration’ (**f**, phase contrast; **g**, TUNEL staining; **h**, neurofilament staining; **i** shows **g** and **h** merged). Surviving axons (long arrows in **f**, **h**, **i**) originate from a position (1) in the explant showing little or no TUNEL staining (**g**, **i**), in contrast to the TUNEL-positive areas in immediately adjacent regions, which are sites of origin of axons that have disintegrated (such as axons shown by short arrows in **f** originating from site 2), again indicating that the factor must be acting on axons rather than cell bodies. Small asterisk indicates a cell-sparse region in this focal plane (as visualized with neurofilament stain **h**) and with DAPI (not shown). Scale bar: 100 μm .

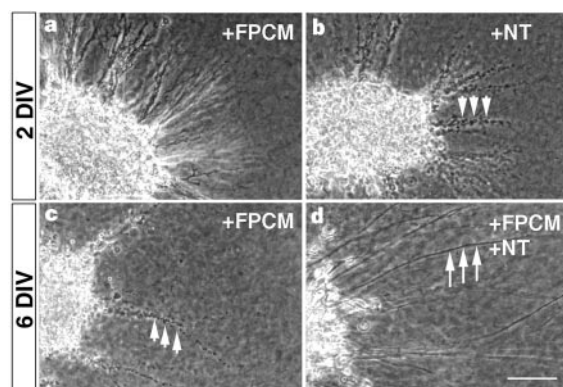


Figure 6 Commissural neurons acquire dependence on neurotrophins (NT) at a developmental stage equivalent to E17. E13 explants were cultured with 30 ng ml^{-1} netrin-1 to elicit commissural axon outgrowth. **a, c**, Adding FPCM after 24 h of culture prevents death of axon bundles at 2 DIV ($n = 20$ explants) (**a**), and if added daily prevents death until 4 DIV, with death being evident at 6 DIV (small arrows show an example of disintegrating axons) ($n = 18$ explants) (**c**). **b, d**, Adding neurotrophins (NGF, 50 ng ml^{-1} ; BDNF, 50 ng ml^{-1} ; NT-3, 25 ng ml^{-1} ; NT-4/5, 20 ng ml^{-1}) instead of FPCM at 24 h does not prevent the death that occurs between 24 and 48 h (all 20 explants) (**b**), but addition of neurotrophins at 4 DIV to explants cultured with FPCM (added daily) suppresses the cell death seen at 6 DIV with FPCM alone (long arrows indicate healthy axons) (22 of 24 explants) (**d**). Similar survival-promoting effects are seen at 6 DIV with a combination of just NGF, NT3 and BDNF (data not shown). Scale bar: 50 μm .

cells at around the time (E13–E14) that they start to contact the floor plate, that this dependence is maintained during and possibly after their journey alongside the floor plate, and that they acquire additional trophic dependencies as they mature further and reach their targets. The findings of a steep gradient of trophic activity along the dorsoventral axis of the spinal cord and of a local trophic effect of floor plate cells on the axons of commissural neurons *in vitro* indicate that the effect of the floor plate is mediated by a trophic signal sensed by the axons close to the floor plate, although we cannot exclude that some trophic support might be provided at the cell bodies of these neurons by a trophic factor either produced in the dorsal spinal cord or diffusing there from floor plate cells. Our experiments do not address whether, earlier in their trajectory, immediately after being born but before reaching the floor plate, the commissural neurons are also dependent on trophic support provided locally by cells within the DSC, nor do we know whether acquisition of novel trophic dependences is a cell-autonomous phenomenon dependent on an intrinsic clock or is regulated by extrinsic signals to which the neurons are exposed.

The fact that many axons must grow for a considerable period to reach their distant targets has prompted previous investigators to wonder whether “during this period . . . the cells contain within themselves all that they need for their growth and survival, or whether these cells depend on a continuous supply of . . . exogenous trophic input from their immediate environment”¹¹. The possible existence of intermediate target-derived trophic support for extending axons has been suggested by recent studies of sensory neurons in the peripheral nervous system¹². Although sensory neurons appear to be independent of trophic support for one or more days after they differentiate^{13,14}, evidence that spinal sensory neurons en route to their targets obtain trophic support from intermediate target-derived NT-3 is suggested by the enhanced death of these neurons in NT-3 knock-out mice at a time (E11) when their axons are beginning to reach the vicinity of their target fields¹⁵. This suggestion is supported by the fact that NT-3 is available at least to pioneering sensory axons as they migrate¹⁵, but is complicated by the facts that NT-3 may also be available to sensory cell bodies¹⁵ and that the precise timing of death^{15,16} is potentially consistent with a role for NT-3 derived from the vicinity of their differentiating target fields. Studies of premature cell death in the trigeminal ganglia of mice deficient in various neurotrophins or their receptors similarly suggest an early dependence of these neurons on neurotrophins, but it remains unclear whether the dependence is initiated before or soon after axons reach the vicinity of their differentiating target fields (refs 17–19 and L. Reichardt, personal communication). Although additional studies are required to clarify the role of intermediate target-derived trophic support for extending axons in the peripheral nervous system, our findings nonetheless provide evidence for the operation of this mechanism in the CNS. The dependence of neurons on trophic support from their final targets has been suggested to provide a means for eliminating misprojecting neurons²⁰. However, an earlier dependence on trophic support provided *en passant* by intermediate targets, as documented here, potentially provides a much more effective means of eliminating such neurons as soon as they make their errors, thus denying them the opportunity to wander and by chance to find alternative trophic support and thereby to establish aberrant neuronal circuits. □

Methods

Explant culture and histochemistry.

E11–E14 DSC explants and various targets were dissected and embedded in three-dimensional collagen gels and cultured in supplemented Ham's F12 medium⁴ containing 5% heat-inactivated horse serum (HIHS) (GIBCO) as described^{4,5,7}. FPCM was prepared by culturing floor plate tissue dissected from four E13 embryos in 100 µl of medium in cell culture inserts (0.4 µm pore size, Fisher) for 1 day before transferring the inserts to wells containing DSC explants. In experiments to examine the effect of neurotrophins at later stages, only the medium from the cell culture insert was transferred to the wells containing dorsal explants. Whole-mount immunohistochemistry was performed on E11 and E13 rat

dorsal explants using antibodies to neurofilament as described⁶. TUNEL assay was performed using the *In Situ* Cell Death Detection kit (Boehringer Mannheim). The protein concentration of tissue fragments (Fig. 4 legend) was measured by Bradford assay (BioRad) after tissues were dissolved in 10 mM Tris-HCl (pH 7.4), 0.1% SDS and boiled for 10 min.

Assay of defined factors.

Factors were added to E13 DSC explants grown with 30 ng ml⁻¹ netrin-1 after 24 h. We assayed axon disintegration after an additional 16–24 h of incubation. The factors tested were: BDNF; bone morphogenic proteins 1 (BMP1) and 2; bone morphogenic protein 7/osteogenic protein 1; ciliary neurotrophic factor; basic fibroblast growth factor; glial-derived neurotrophic factor; glial growth factor 2; heregulin; hepatocyte growth factor; insulin-like growth factor; interleukin 3; interleukin 6; interleukin 7; leukemia inhibitory factor; netrin-1; NGF; NT-3; neurotrophin 4/5; neurturin; platelet-derived growth factor AA; platelet-derived growth factor BB; persephin; Steel Factor/Kit-Ligand; transforming growth factor-α; transforming growth factor-β; and transforming growth factor-β2. Each factor was tested alone, or in the following combinations: all neurotrophins as a group, or all 27 factors, or all 27 factors in combination with 5 µM forskolin and 0.1 mM isobutylmethylxanthine (Sigma) (the latter potentiate the survival effects of various growth factors on some neuronal populations)²¹. All factors were from human and were tested at 10 ng ml⁻¹ and 40 ng ml⁻¹, except netrin-1 which is from chick and tested at 30 ng ml⁻¹ and BMP1, tested up to 1 µg ml⁻¹.

Received 7 April; accepted 2 July 1999.

- Oppenheim, R. W. Neurotrophic survival molecules for motoneurons: an embarrassment of riches. *Neuron* **17**, 195–197 (1996).
- Pettmann, B. & Henderson, C. E. Neuronal cell death. *Neuron* **20**, 633–647 (1998).
- Altman, J. & Bayer, S. A. The development of the rat spinal cord. *Adv. Anat. Embryol. Cell Biol.* **85**, 1–164 (1984).
- Tessier-Lavigne, M., Placzek, M., Lumsden, A. G., Dodd, J. & Jessell, T. M. Chemotropic guidance of developing axons in the mammalian central nervous system. *Nature* **336**, 775–778 (1988).
- Placzek, M., Tessier-Lavigne, M., Jessell, T. & Dodd, J. Orientation of commissural axons *in vitro* in response to a floor plate-derived chemoattractant. *Development* **110**, 19–30 (1990).
- Kennedy, T. E., Serafini, T., de la Torre, J. R. & Tessier-Lavigne, M. Netrins are diffusible chemotropic factors for commissural axons in the embryonic spinal cord. *Cell* **78**, 425–435 (1994).
- Serafini, T. *et al.* The netrins define a family of axon outgrowth-promoting proteins homologous to *C. elegans* UNC-6. *Cell* **78**, 409–424 (1994).
- Serafini, T. *et al.* Netrin-1 is required for commissural axon guidance in the developing vertebrate nervous system. *Cell* **87**, 1001–1014 (1996).
- Keino-Masu, K. *et al.* Deleted in Colorectal Cancer (DCC) encodes a netrin receptor. *Cell* **87**, 175–185 (1996).
- Compenot, R. B. Development of sympathetic neurons in compartmentalized culture. I. Local control of neurite growth by nerve growth factor. *Dev. Biol.* **93**, 1–12 (1982).
- Cowan, W. in *Symposia Celebrating the Harvard Medical School Bicentennial* (ed. Isselbacher, K.) 643–688 (Wiley, New York, 1984).
- Davies, A. M. Neurotrophin switching: where does it stand? *Curr. Opin. Neurobiol.* **7**, 110–118 (1997).
- Buchman, V. & Davies, A. Different neurotrophins are expressed and act in a developmental sequence to promote the survival of embryonic sensory neurons. *Development* **118**, 989–1001 (1993).
- Wright, E., Vogel, V. K. & Davies, A. Neurotrophic factors promote the maturation of developing sensory neurons before they become dependent on these factors for survival. *Neuron* **9**, 139–150 (1992).
- Farinas, I., Yoshida, C., Backus, C. & Reichardt, L. Lack of neurotrophin-3 results in death of spinal sensory neurons and premature differentiation of their precursors. *Neuron* **17**, 1065–1078 (1996).
- Liebl, D., Tessarollo, L., Palko, M. & Parada, L. Absence of sensory neurons before target innervation in brain-derived neurotrophic factor, neurotrophin-3 and TrkC-deficient embryonic mice. *J. Neurosci.* **17**, 9113–9121 (1997).
- Wilkinson, G. A., Farinas, I., Backus, C., Yoshida, C. K. & Reichardt, L. F. Neurotrophin-3 is a survival factor *in vivo* for early mouse trigeminal neurons. *J. Neurosci.* **16**, 7661–7669 (1996).
- Pinon, L., Minichiello, L., Klein, R. & Davies, A. Timing of neuronal death in trkA, trkB and trkC mutant embryos reveals developmental changes in sensory neuron dependence on Trk signalling. *Neuron* **122**, 3255–3261 (1996).
- Huang, E. *et al.* Expression of Trk receptors in the developing mouse trigeminal ganglion: *in vivo* evidence for NT-3 activation TrkA and TrkB in addition to TrkC. *Development* **126**, 2191–2203 (1999).
- Purves, D. & Lichtman, J. W. *Principles of Neural Development* (Sinauer, Sunderland, Massachusetts, 1985).
- Meyer-Franke, A., Kaplan, M., Pfriger, F. & Barres, B. Characterization of the signaling interactions that promote the survival and growth of developing retinal ganglion cells in culture. *Neuron* **15**, 805–819 (1995).

Acknowledgements

We thank B. Barres for providing most of the factors tested, Amgen for the Steel Factor, D. Greenspan for BMP-1 protein, and R. O'Connor, J. Flanagan, M. W. Cowan, L. Reichardt, K. Brose, M. Gallo and P. Leighton for helpful discussions and/or comments on the manuscript. Supported by a fellowship from the American Cancer Society to H.W. and by a grant from the National Institutes of Health to M.T.-L. M.T.-L. is an Investigator of the Howard Hughes Medical Institute.

Correspondence and requests for materials should be addressed to M.T.-L. (e-mail: marctl@itsa.ucsf.edu).

Persistent patterns in transient chaotic fluid mixing

D. Rothstein*, E. Henry* & J. P. Gollub*†

* Department of Physics, Haverford College, Haverford, Pennsylvania 19041, USA

† Department of Physics, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA

Chaotic advection^{1–3} of a fluid can cause an initially inhomogeneous impurity (a passive scalar field) to develop complex spatial structure as the elements of the fluid are stretched and folded, even if the velocity field is periodic in time. The effect of chaotic advection on the transient mixing of impurities—the approach to homogeneity—has been explored theoretically and numerically^{4–8}. A particularly intriguing prediction is the development of persistent spatial patterns, whose amplitude (contrast) decays slowly with time but without change of form. Here we investigate these phenomena using an electromagnetically driven two-dimensional fluid layer in which one half is initially labelled by a fluorescent dye (the passive scalar). We observe the formation of structurally invariant but slowly decaying mixing patterns, and we show how the various statistical properties that characterize the dye concentration field evolve with time as mixing proceeds through many cycles. These results show quantitatively how advective stretching of the fluid elements and molecular diffusion work together to produce mixing of the impurity. We contrast the behaviour of time-periodic flows and identically forced but weakly turbulent flows at lower viscosity, where mixing is much more efficient.

Density stratification and magnetic forcing are used to create a two-dimensional flow with good accuracy. A layer of glycerol–water–salt solution 3 mm deep carries a horizontal electric current in the presence of an array of strong permanent magnets just below the fluid. An ordered square array with alternating polarity, and a spatially disordered array (with magnets having random locations

and polarity), are used to create both regular and disordered cellular flows by means of Lorenz forces.

The fluid of interest is a dilute fluorescent dye, contained in a separate 1 mm upper layer of glycerol–water (37% by weight). It is less dense than the glycerol–water–salt solution described above, and the layers remain separate for the duration of an experiment, typically 10 min. The dye diffuses slowly because the Schmidt number (the ratio of kinematic viscosity to dye diffusivity) is about 2,000. The upper layer, which contains no salt, is driven by friction at its lower boundary. This arrangement provides a nearly two-dimensional flow in the upper layer^{9,10}. Residual vertical flows are at least an order of magnitude weaker than the horizontal circulation. Only half the upper layer is initially labelled with the dye.

The use of glycerol (32% by volume) suppresses instabilities leading to turbulence by increasing the viscosity to $3.3 \times 10^{-6} \text{ m}^2 \text{ s}^{-1}$, and thus keeping the Reynolds number well below 100. As a result, when the current is steady (and of the order of 60 mA), the velocity field $\mathbf{v}(\mathbf{r})$ is time-independent. When the current is oscillatory, $\mathbf{v}(\mathbf{r})$ is found to be time-periodic. The fluid cell is placed in the feedback loop of an operational amplifier circuit so that the current accurately follows the applied voltage waveform. For time-periodic square or sine wave forcing there is no mean flow. When the glycerol is omitted, on the other hand, two-dimensional weakly turbulent flows can be created¹¹.

The fluid is illuminated near 365 nm, and the fluorescent response, imaged by a cooled 12 bit CCD camera operating at 512×512 pixel resolution, is accurately linear in the local concentration under the experimental conditions¹⁰. The camera is triggered to sample the dye distribution every period or half period.

An example of the mixing process is shown in Fig. 1 for the case of time-periodic forcing at 100 mHz after 20 periods, using a spatially periodic magnet array. The dye solution, initially confined to the right half of the image, is progressively stretched and folded,



Figure 1 Snapshot of the concentration field for transient magnetically forced chaotic mixing. Fluid labelled with fluorescein was initially confined to the right half of the 20×24 cm cell; the imaged region is 9.4×9.4 cm. The light intensity under ultraviolet illumination is shown after 20 periods of forcing by a time periodic electric current at 100 mHz, in the presence of a spatially periodic magnet array. The repetitive stretching and folding that is characteristic of chaotic advection is evident.

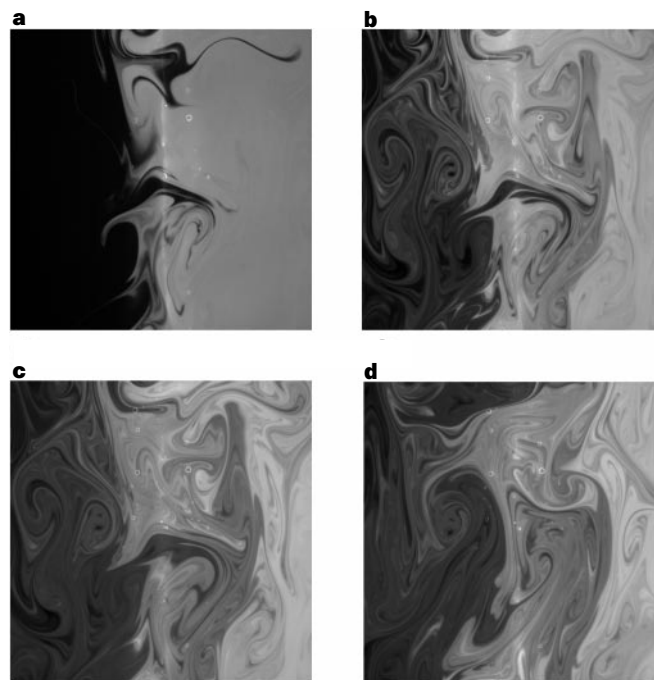


Figure 2 Attainment of a persistent mixing pattern using a disordered array of forcing magnets. The impurity field develops a complex structure after about 10 periods that repeats periodically, except for a gradual loss of contrast. **a–c**, $t = 2, 20$ and 50 periods, respectively. **d**, Sampling out of phase at 50.5 periods shows that the tracer distribution changes substantially within each cycle. The forcing frequency is 70 mHz.

intruding into the left half, while unlabelled fluid penetrates the right half. The underlying forcing vortices are evident from the dye pattern, and the striations are typical of chaotic advection.

The case of time-periodic forcing by a random magnet array is shown as a function of time in Fig. 2, by means of images obtained at constant phase relative to the forcing after 2, 20 and 50 cycles. Remarkably, after about 10 cycles, a complex spatial pattern is reached that subsequently recurs once per period. This may be seen qualitatively by comparing Fig. 2b and c, which are nearly identical, including the fine detail, except for a reduction in contrast over time. The final panel shows the situation one half cycle later at 50.5 cycles, and indicates that the impurity distribution is quite different (though statistically similar) at other phases relative to the forcing.

To demonstrate the approach to a persistent pattern quantitatively, we show (Fig. 3a) the cross-correlation coefficient of images obtained 4 cycles apart, as a function of time from the start of the mixing. This quantity increases rapidly from a low value to about 0.98, showing that images obtained 4 cycles apart are nearly identical after about 10 periods of mixing. A similar result is obtained for images spaced 10 cycles apart. Another way to test for a persistent but slowly decaying pattern is to fit the intensity data to a product of separate space and time functions: $I(r, t) = f(r)g(t) + C$, where the constant C is the intensity after complete mixing. This can be done with a r.m.s. variation of less than 1% for $10 < t < 70$ periods. The resulting function $f(r)$ is similar to Fig. 2c, while the contrast $g(t)$ decays smoothly to zero, leaving a uniform intensity. We find that similar fits can be obtained for the spatially

periodic forcing of Fig. 1.

Although the spatial structure approaches an asymptotic form, mixing continues, owing to stretching of fluid elements and diffusion. This may be seen by monitoring the intensity probability distribution function (PDF) and the decay of its normalized variance $\langle(\delta I)^2\rangle$, which is a measure of "unmixedness"¹² or contrast, shown in Fig. 3b. The variance of the PDF shows a pronounced tail, and is approximately described by a power law of the form $bt^{-0.69 \pm 0.01}$ for $t > 10$, where b is a constant. However, for longer runs, an approximately exponential decay is eventually attained. Also, in Fig. 3b, we show the decay of the variance when the viscosity is reduced so that the flow becomes weakly turbulent at higher Reynolds number, with all forcing parameters the same. The decay is then much faster and is well described as an exponential in time of the form $c \exp[-(0.7 \pm 0.1)t]$. However, there is no well-defined asymptotic pattern for the turbulent case; the concentration field pattern continues to fluctuate strongly. The fluorescent intensity PDFs are also quite different for the two cases: the turbulent flow PDF approaches a gaussian for $t \geq 50$, while that for the chaotic flow is complex and broader.

We have tested for the approach to a persistent pattern under a variety of other circumstances: (1) if the forcing frequency is increased, the approach to the asymptotic state requires more cycles because the shorter fluid trajectories before reversal tend to reduce the effective Lyapunov exponent; (2) if the forcing current contains a d.c. component so that there is a mean flow, the results are not substantially changed; (3) we find generally similar (but

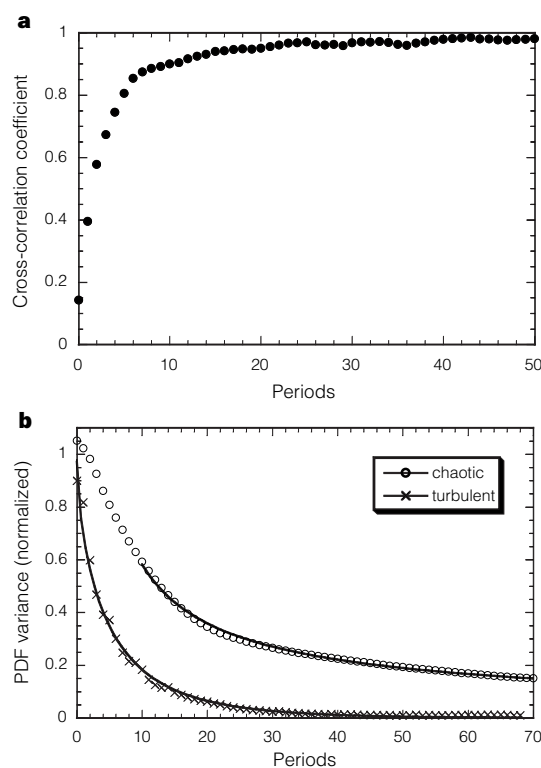


Figure 3 Quantitative characterization of the persistent mixing pattern. **a**, Cross-correlation coefficient between images separated by four periods, as a function of time from the start of mixing, for the chaotic case of Fig. 2. (The intensity is measured with respect to the average along a vertical line, in order to remove spurious correlation due only to the background dye gradient.) The asymptotic value of 0.98 implies that the concentration field returns to almost the same state once per cycle. **b**, Variance $\langle(\delta I)^2\rangle$ of the fluorescent intensity probability distribution function (PDF) versus time, for chaotic mixing as in Fig. 2, and for mixing by weakly turbulent flow at lower viscosity using 70 mHz forcing. The Reynolds number is about ten times larger in the turbulent case. Fits to the data are described in the text.

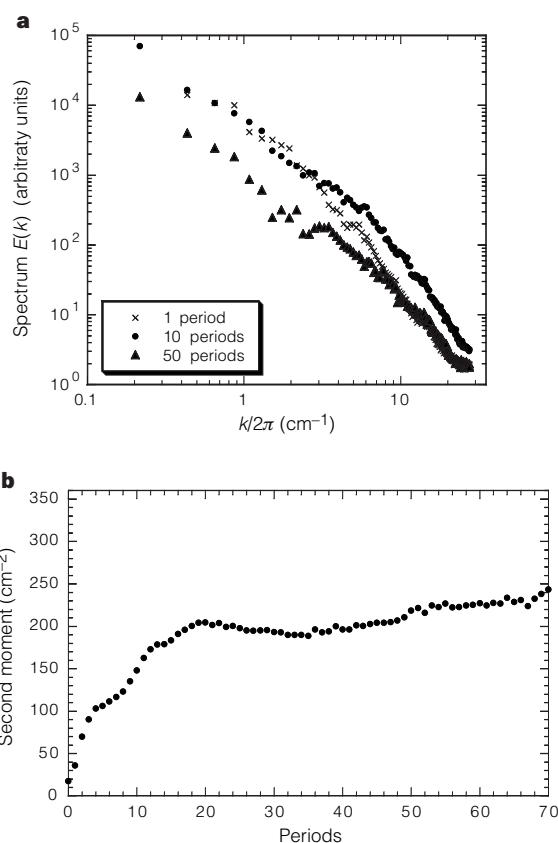


Figure 4 Spatial power spectra and moments for transient mixing. **a**, One-dimensional power spectrum $E(k)$ at several times, for the chaotic advection of Fig. 2. The spectrum widens over the first few periods, as a result of stretching and folding. Later, the shape remains roughly fixed but the overall amplitude decreases with time at all k values. **b**, Second spectral moment as a function of time, showing saturation after about 10 periods.

somewhat slower) mixing for spatially periodic forcing; (4) for temporally non-periodic forcing, there are no recurrences.

The one-dimensional spatial power spectrum $E(k)$ computed in the standard way¹⁰ provides a useful diagnostic for the development of fine structure in the concentration field. An example is shown in Fig. 4a for time-periodic forcing using the spatially disordered magnet array (as in Fig. 2), at several different times. For $l < t < 10$, the high frequency tail of $E(k)$ rises, as stretching and folding produce fine structure. Later, the entire spectrum declines at all frequencies with little further change of shape. The decline at high k is due to diffusion, and the decline at low k is due to the continuing transport of spectral variance to high k by stretching. The spectrum does not follow a well-defined power law (nor is one expected⁶ for transient mixing); it is steeper at higher k . The behaviour of the second spectral moment $\langle k^2 \rangle$, shown in Fig. 4b, increases rapidly at first and then essentially saturates, consistent with the asymptotic structure noted earlier.

The fine structure of the dye pattern may also be investigated by studying the PDF of the magnitude of the concentration gradient. An example is shown in Fig. 5 for the same conditions as in Fig. 2. Gradients have been expressed in units of the standard deviation of the concentration field to compensate for the gradual homogenization. It is striking to see that the gradient PDF reaches an invariant form that is essentially identical at $t = 20$ and 50 periods. The distribution at high gradients is accurately exponential. Such exponential tails for distributions of concentration gradients have been noted in several theoretical studies¹³. When the same experiment is done for the weakly turbulent flow at lower viscosity, the resulting distribution is similar to that of Fig. 5 for $t = 10$, but the shape continues to evolve, approaching a maxwellian distribution (characteristic of independent gaussian components) at longer times. Therefore, the large gradients are attenuated more rapidly than are small gradients for the turbulent flow. Diffusion broadens the striations in two-dimensional turbulence, but not in chaotic mixing.

We find that chaotic mixing, produced by time-periodic cellular flows, leads to a remarkable persistent spatial structure, a complex pattern that recurs periodically while its contrast decays slowly with time. The recurrence is surprising (when first encountered) since additional striations are created during each cycle. The corresponding spatial power spectrum retains its shape as it declines simultaneously at all wavenumbers owing to a delicate balance of two

distinct processes: stretching and diffusion. The probability distribution of the magnitude of the concentration gradient reaches an invariant form (when suitably normalized) with a peak at non-zero gradient and exponential tails. All of these properties show that the recurrent mixing pattern represents a dynamical non-equilibrium state rather than a static condition. Finally, we note that identically forced but weakly turbulent flows at lower viscosity become homogeneous and transport material much more rapidly than the time-periodic flows manifesting chaotic advection. However, turbulence is generally unavailable when mixing in small scale devices is required. □

Received 3 June; accepted 18 August 1999.

1. Aref, H. Stirring by chaotic advection. *J. Fluid Mech.* **143**, 1–21 (1984).
2. Ottino, J. M. Mixing, chaotic advection, and turbulence. *Annu. Rev. Fluid Mech.* **22**, 207–253 (1990).
3. Fountain, G. O., Khakhar, D. V. & Ottino, J. M. Visualization of three-dimensional chaos. *Science* **281**, 683–686 (1998).
4. Pierrehumbert, R. T. Tracer microstructure in the large-eddy dominated regime. *Chaos Solitons Fractals* **4**, 1091–1110 (1994).
5. Antonsen, T. M., Fan, Z., Ott, E. & Garcia-Lopez, E. The role of chaotic orbits in the determination of power spectra of passive scalars. *Phys. Fluids* **8**, 3094–3104 (1996).
6. Antonsen, T. M. & Ott, E. Multifractal power spectra of passive scalars convected by chaotic fluid flows. *Phys. Rev. A* **44**, 851–857 (1991).
7. Muzzio, F. J., Meneveau, C., Swanson, P. D. & Ottino, J. M. Scaling and multifractal properties of mixing in chaotic flows. *Phys. Fluids A* **4**, 1439–1456 (1992).
8. Alvarez, M. M., Muzzio, F. J., Cerbelli, S., Adrover, A. & Giona, M. Self-similar spatiotemporal structure of intermaterial boundaries in chaotic flows. *Phys. Rev. Lett.* **81**, 3395–3398 (1998).
9. Paret, J., Marteau, D., Paireau, O. & Tabeling, P. Are flows electromagnetically forced in thin stratified layers two-dimensional? *Phys. Fluids* **9**, 3102–3104 (1997).
10. Williams, B. S., Marteau, D. & Gollub, J. P. Mixing of a passive scalar in two-dimensional turbulence. *Phys. Fluids* **9**, 2061–2080 (1997).
11. Paret, J. & Tabeling, P. Experimental observation of the two-dimensional inverse energy cascade. *Phys. Rev. Lett.* **79**, 4162–4165 (1997).
12. Miller, P. L. & Dimotakis, P. E. Reynolds number dependence of scalar fluctuations in a high Schmidt number turbulent jet. *Phys. Fluids A* **3**, 1156–1163 (1991).
13. Shraiman, B. I. & Siggia, E. D. Lagrangian path integrals and fluctuations in random flow. *Phys. Rev. E* **49**, 2912–2927 (1994).

Acknowledgements

We thank J. Andersen, J.-C. Geminard, A. Kudrolli, and W. Losert for experimental contributions, and E. Ott and B. Shraiman for discussions. This work was supported by the Condensed Matter Physics Program of the US NSF.

Correspondence and requests for materials should be addressed to J.P.G. (e-mail: jgollub@haverford.edu).

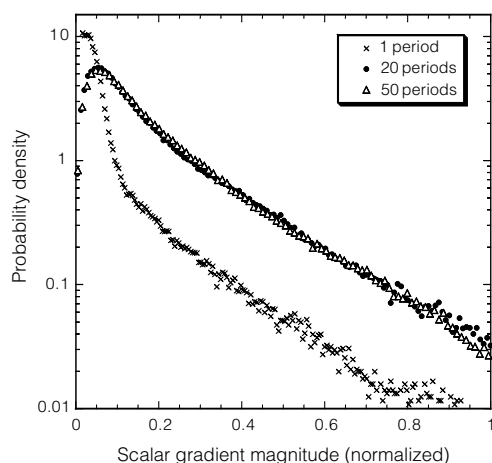


Figure 5 Probability distribution of the magnitude of the concentration gradient for chaotic mixing (Fig. 2). The distribution, here normalized by the standard deviation of the intensity field to compensate for the gradual loss of contrast, reaches an invariant form after about 20 periods, with a peak at finite gradient and an exponential tail.

Sonoluminescence temperatures during multi-bubble cavitation

William B. McNamara III*, Yuri T. Didenko*‡ & Kenneth S. Suslick*

* Department of Chemistry, University of Illinois at Urbana-Champaign, 600 S. Mathews Avenue, Urbana, Illinois 61801, USA

Acoustic cavitation—the formation and implosive collapse of bubbles—occurs when a liquid is exposed to intense sound. Cavitation can produce white noise, sonochemical reactions, erosion of hard materials, rupture of living cells and the emission of light, or sonoluminescence^{1,2}. The concentration of energy during the collapse is enormous: the energy of an emitted photon can exceed the energy density of the sound field by about twelve orders of magnitude³, and it has long been predicted that the interior bubble temperature reaches thousands of degrees Kelvin during collapse. But experimental measurements^{4,5} of

‡ Permanent address: Pacific Oceanological Institute, Vladivostok 690061, Russia.

conditions inside cavitating bubbles are scarce, and there have been no studies of interior temperature as a function of experimental parameters. Here we use multi-bubble sonoluminescence from excited states of metal atoms as a spectroscopic probe of temperatures inside cavitating bubbles. The intense atomic emission allows us to change the properties of the gas–vapour mixture within the bubble, and thus vary the effective emission temperature for multi-bubble sonoluminescence from 5,100 to 2,300 K. We observe emission temperatures that are in accord with those expected from compressional heating during cavitation.

We previously discovered that sonication of solutions of $\text{Fe}(\text{CO})_5$, $\text{Cr}(\text{CO})_6$ and $\text{Mo}(\text{CO})_6$ produced multi-bubble sonoluminescence (MBSL) spectra from excited states of iron, chromium, and molybdenum atoms⁶. Improved spectral resolution now enables us to use these emissions to determine effective temperatures for the emitting excited states. As an example, calculated and observed spectra for Cr under Ar at 20 kHz are given in Fig. 1; the synthetic spectra are generated using the theory of atomic emission⁷. Analogous spectra were also collected for Mo and Fe in order to test the validity of this method. The relative integrated intensities (I_1/I_2) of two atomic lines emitted from the different excited states of the same metal atom are given by equation (1):

$$\frac{I_1}{I_2} = \frac{g_1 A_1 \lambda_1}{g_2 A_2 \lambda_2} \exp[(E_1 - E_2)/kT_e] \quad (1)$$

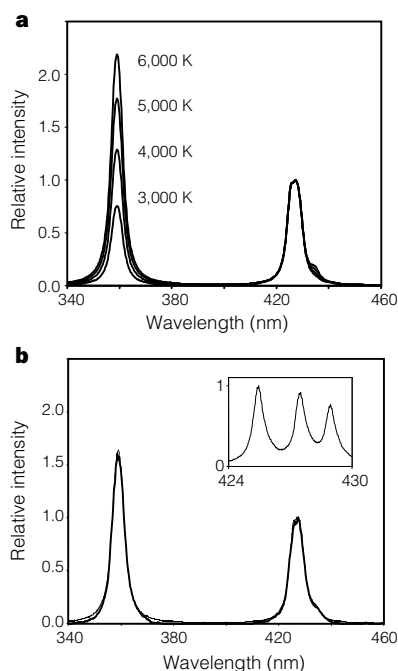


Figure 1 Multi-bubble sonoluminescence (MBSL) emission from excited states of Cr atoms. **a**, Calculated spectra as a function of temperature. **b**, The observed emission spectrum compared to the best-fit calculated spectrum (4,700 K). Inset, observed spectrum at higher resolution, which resolves the individual atomic emission lines. Sonoluminescence was generated from 2.5 mM solutions of metal carbonyls (0.25 mM for $\text{Mo}(\text{CO})_6$ due to solution absorbance) in poly(dimethylsiloxane) (Dow 200 silicone oil, 100 centistokes viscosity) irradiated at $\sim 90 \text{ W cm}^{-2}$ at 20 kHz (Heat Systems Mysonix 375, with 0.5 inch Ti horn), and spectra collected with an 0.5 m spectrograph (Acton Research 505F with a Princeton Instruments IRY 512N diode array). The solutions were continuously sparged with the appropriate gas during sonication, and each spectrum is the average of at least 15 10-second collections. All spectra were corrected for absorbance by the solution and for the response of the optical system. Imaging of the MBSL shows that the emission comes from a dense cloud of emitting bubbles rather than from a few exceptional ones.

where g is the degeneracy of the electronic state, A is the Einstein transition probability for the electronic transition, λ is the wavelength of the emitted line, E is the energy of the electronic state, and T_e is the electronic temperature. The transition probabilities, degeneracies, wavelengths, and excited-state energies were taken from the literature^{8–10}. This technique has been well established for flame and plasma diagnostics^{11,12}, and is proving to be a versatile method of determining the effective temperature of cavitation bubble collapse under a variety of experimental conditions.

The temperatures from three different metal atomic line emissions yielded very consistent results. High-resolution spectra of the MBSL emission from Cr and Mo atoms were collected from solutions of the hexacarbonyl complexes of these metals (Fig. 1 inset). The relative intensity of each line within a given multiplet was measured from these spectra, and the relative intensity of each multiplet was then determined from lower-resolution spectra. This process allowed for a line by line comparison of relative intensities for these two metals across the entire spectral window. The temperatures were then determined by the use of atomic Boltzmann plots¹³: Cr and Mo atom emission temperatures were respectively $4,700 \pm 300 \text{ K}$ and $4,800 \pm 400 \text{ K}$, under Ar at 20 kHz. The linearity of these plots also indicated that the emission is thermally equilibrated. The high resolution MBSL spectra of Fe emission are too complex to allow for the determination of individual line intensities, and the experimental spectra were therefore compared to synthetic spectra generated using equation (1), with an excellent fit to data found for a temperature of $5,100 \pm 300 \text{ K}$.

We previously used two different techniques to measure the temperature of cavitation in Ar-saturated liquids. The first made use of sonochemical kinetics for a comparative rate thermometry analysis of the dissociation of CO from various metal carbonyls⁵. This technique is independent of sonoluminescence, and yielded an effective reaction temperature of $5,200 \pm 500 \text{ K}$ for the gas phase of cavitation under Ar. The second technique applied spectral analysis to C_2 emission bands in the MBSL from the sonication of silicone oil under Ar, giving an effective emission temperature of $5,075 \pm 150 \text{ K}$ (ref. 4). The temperatures reported here are in good agreement with these previously determined values. We note that the previous MBSL emission spectra of C_2 under similar conditions also appeared to be thermally equilibrated^{4,14}.

These metal-atom emissions are the most intense sonoluminescence yet observed. This allows us to study MBSL under conditions previously experimentally inaccessible. Multi-bubble sonoluminescence is generally believed to arise from compressional heating of the bubble contents during the final stage of bubble collapse. The intrabubble temperature should therefore depend on the properties

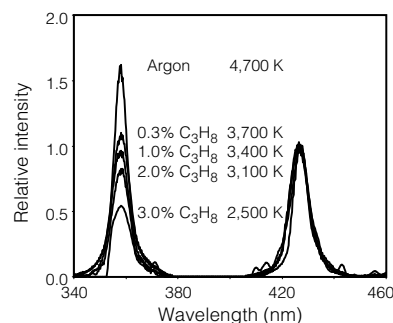


Figure 2 Effect of polyatomic gases on the observed emission temperatures from MBSL. The MBSL spectra are from solutions of $\text{Cr}(\text{CO})_6$ in silicone oil, obtained as in Fig. 1. From top to bottom, the sparge gas was Ar mixed with 0.0, 0.3, 1.0, 2.0, and 3.0 vol % propane, yielding respective emission temperatures of 4,700 K, 3,700 K, 3,400 K, 3,100 K and 2,500 K.

of the gas and vapour mixture within the bubble, particularly the polytropic ratio ($\gamma = C_p/C_v$, ratio of heat capacities) and the thermal conductivity. The effect of each of these parameters was examined by collecting MBSL spectra from several liquids containing a variety of dissolved gases.

If emission results from compressional heating, the effective temperature and pressure reached during cavitation will drop as γ decreases (for a given compression ratio). This prediction was tested by collecting MBSL spectra from $\text{Cr}(\text{CO})_6$ solutions that were saturated with Ar doped with low concentrations of various gaseous hydrocarbons (methane, ethane, ethylene and propane). These polyatomic molecules have much higher heat capacities than does Ar, particularly at temperatures sufficient to populate their excited-state vibrational modes. Note that the thermal conductivities of these mixtures vary by less than a factor of two relative to argon.

The observed emission temperature does indeed decrease substantially as the polyatomic gas content is increased. For example, as Ar is doped with increasing amounts (up to 3%) of propane (Fig. 2), the emission temperature progressively decreases from 4,700 K to 2,500 K. The low resolution of the Cr spectra was necessitated by the decreasing intensity of emission at the higher concentrations of hydrocarbons. The measured emission temperatures for Ar mixtures with various hydrocarbons (methane, ethane, ethylene and propane) are given in Fig. 3. In all instances, the temperature of MBSL decreases as the percentage of a given hydrocarbon is increased (that is, as the polytropic ratio of the gas within the bubble is decreased). Similar results were observed for Fe with methane.

The very simplest model for cavitation assumes adiabatic compression during bubble collapse^{2,15}: a gas with a polytropic ratio γ , compressed from an initial radius R_i to a final radius R_f , with an initial temperature T_i , has a final temperature T_f according to equation (2):

$$T_f = T_i \left(\frac{R_i}{R_f} \right)^{3(\gamma-1)} \quad (2)$$

This equation includes no dynamics of bubble motion, and neglects all effects from thermal transport, solvent vapour pressure, and chemical reactions within the bubble. Furthermore, equation (2) assumes a temperature-independent γ ; however, γ decreases substantially as excited vibrational states become fully populated and it decreases even more if chemical reactions occur. Nonetheless, as shown in Fig. 3, the data are fitted reasonably well by equation (2) for a compression ratio of 3.65. We calculated the temperatures and

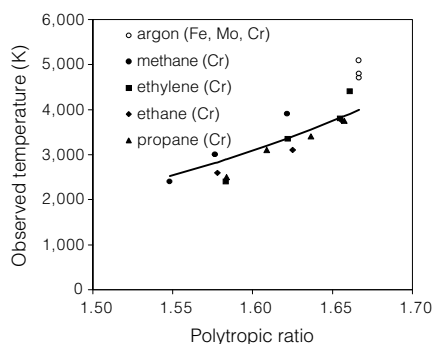


Figure 3 Observed emission temperatures versus gas-mixture polytropic ratios at 298 K. The polytropic ratio of heat capacities, γ , is given by C_p/C_v . The solid line is the best fit to simple adiabatic compression of a bubble with a radius ratio of 3.65, from equation (2); as γ is not temperature independent, equation (2) provides only an approximate model. The open circles give the observed emission temperatures from Fe, Mo, and Cr atom MBSL under argon; all other data are from Cr atom emission under given gases.

pressures during bubble collapse by finite difference using a temperature dependent γ , and can fit our data to within experimental error. Given the lack of dynamics and the failure to incorporate chemical reactions (whose endothermicity is not negligible compared to the compressional heating), more detailed modelling¹⁶ will be required to give meaningful results.

The polytropic ratio of the bubble interior can also be varied by changing the identity and vapour pressure of the solvent. The solvent vapours behave in the same manner as the gaseous hydrocarbons described above: an increase in the solvent vapour pressure is equivalent to increasing the percentage of hydrocarbon gas and leads to similar drops in temperature. Both MBSL intensities^{17–19} and sonochemical rates^{20,21} decrease dramatically with solvent vapour pressure increases. We collected MBSL spectra at different bulk solution temperatures from solutions of $\text{Cr}(\text{CO})_6$ under Ar in octanol, a low volatility liquid whose vapour pressure as a function of temperature is well documented²². The results in Fig. 4 are consistent with compressional heating of the bubble contents: the temperature of MBSL decreases as the vapour pressure increases. Similar results were observed using other long-chain alcohols and alkanes, and using helium instead of argon.

Several reports have shown that the total intensity of MBSL from solutions saturated with noble gases decreases as the thermal conductivity of the noble gas increases. This has generally been attributed to thermal transport during cavitation: compressional heating combined with limited thermal transport in a “quasi-adiabatic”

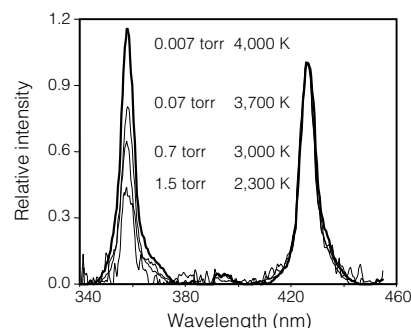


Figure 4 Effect of solvent vapour pressure on the observed emission temperatures from MBSL. The MBSL spectra are from solutions of $\text{Cr}(\text{CO})_6$ in octanol, obtained as in Fig. 1. From top to bottom, the vapour pressure was 0.007, 0.07 and 1.5 torr, yielding emission temperatures of 4,000 K, 3,700 K, 3,000 K and 2,300 K, respectively.

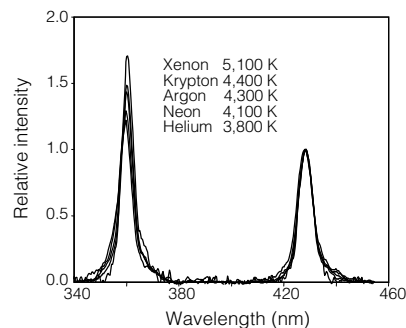


Figure 5 Effect of dissolved noble gases on the observed emission temperatures from MBSL. The MBSL spectra are from solutions of $\text{Cr}(\text{CO})_6$ in octanol, obtained as in Fig. 1. From top to bottom, the dissolved gas was Xe, Kr, Ar, Ne and He, with respective emission temperatures of 5,100 K, 4,400 K, 4,300 K, 4,100 K and 3,800 K.

collapse^{16,23–25}. The rate at which energy is transferred out of the bubble is expected to increase as the thermal conductivity of the gas within the bubble increases. Bubbles containing helium, therefore, should be cooler than bubbles containing argon or xenon, although the predicted extent of this effect is debatable²⁴. In order to examine the importance of thermal conduction, MBSL spectra were obtained from solutions of Cr(CO)₆ in octanol saturated with different noble gases. The results in Fig. 5 show that the temperature of MBSL decreases as the thermal conductivity of the gas within the bubble increases. Similar results were observed using dodecane.

We note that the thermal conductivities of the noble gases used here span a factor of thirty, while the thermal conductivities of the argon and hydrocarbon mixtures differ by only a factor of two, therefore, the temperature changes observed upon the addition of hydrocarbons to argon are not due to changes in thermal conductivity.

We have shown that the observed emission temperature within imploding cavitation bubbles depends on both the polytropic ratio and the thermal conductivity of the bubble contents. The processes occurring within the bubble, however, are more complex than can be described by the simple physics of compressional heating. A cavitation bubble functions as a small chemical reactor, and so the identity of the bubble contents will change during bubble collapse. The sonolysis of organic liquids generates H₂ and C₂H₂ as well as other small hydrocarbons²⁰, and similar processes will occur during the sonolysis of silicone oil. Evidence for this is seen in the MBSL spectra: the spectra of all the systems described here contain features attributable to excited states of C₂ and CH (refs 4, 18, 26), and we also find that sonication of silicone oil leads to intense emission from excited Si atoms. These chemical processes can absorb large amounts of energy during the compression. Furthermore, if sufficient dissociation occurs, the sonolysis products (mostly diatomics and polyatomics) will substantially lower the polytropic ratio and increase the pressure inside the bubble. This makes it increasingly difficult to further compress and heat the bubble²⁷. These factors will limit the temperatures achievable by compressional heating within cavitation bubbles. □

Received 4 May; accepted 23 August 1999.

- Suslick, K. S. & Crum, L. A. in *Handbook of Acoustics* (ed. Crocker, M. J.) 243–253 (Wiley-Interscience, New York, 1998).
- Knapp, R. T., Daily, J. W. & Hammit, F. G. *Cavitation* (McGraw-Hill, New York, 1970).
- Barber, B. P. & Putterman, S. J. Observation of synchronous picosecond sonoluminescence. *Nature* **352**, 318–320 (1991).
- Flint, E. B. & Suslick, K. S. The temperature of cavitation. *Science* **253**, 1325–1326 (1991).
- Suslick, K. S., Hammerton, D. A. & Cline, R. E. Jr Sonochemical hot spot. *J. Am. Chem. Soc.* **108**, 5641–5642 (1986).
- Suslick, K. S., Flint, E. B., Grinstaff, M. W. & Kemper, K. A. Sonoluminescence from metal carbonyls. *J. Phys. Chem.* **97**, 3098–3099 (1993).
- Alkemande, C. T. J., Hollander, T., Snelleman, W. & Zeegers, P. J. T. *Metal Vapours in Flames* (Pergamon, New York, 1982).
- Wiese, W. L., Fuhr, J. R. & Martin, G. A. Atomic transition probabilities: scandium through manganese. *J. Phys. Chem. Ref. Data* **17**, 311–327 (1988).
- Wiese, W. L., Fuhr, J. R. & Martin, G. A. Atomic transition probabilities: iron through nickel. *J. Phys. Chem. Ref. Data* **14**, 13–67 (1988).
- Whaling, W. & Brault, J. W. Comprehensive transition probabilities in molybdenum I. *Phys. Scr.* **38**, 707–715 (1988).
- Gaydon, A. G. *The Spectroscopy of Flames* 2nd edn (Wiley, New York, 1974).
- Cabannes, F. & Chapelle, J. in *Reactions Under Plasma Conditions* 1st edn Vol. 1 (ed. Venugopalan, M.) 367–470 (Wiley-Interscience, New York, 1971).
- Reif, I., Fassel, V. A. & Kniseley, R. N. Spectroscopic flame measurements and their physical significance I: theoretical concepts. *Spectrochim. Acta B* **28**, 105–123 (1973).
- Jeffries, J. B., Copeland, R. A., Suslick, K. S. & Flint, E. B. Thermal equilibration during cavitation. *Science* **256**, 248 (1992).
- Noltingk, B. E. & Neppiras, E. Cavitation produced by ultrasonics. *Proc. Phys. Soc. B* **63**, 674–685 (1950).
- Kamath, V., Prosperetti, A. & Egolfopoulos, F. N. A theoretical study of sonoluminescence. *J. Acoust. Soc. Am.* **94**, 248–260 (1993).
- Jarman, P. Measurements of sonoluminescence from pure liquids and some aqueous solutions. *Proc. Phys. Soc.* **73**, 628–640 (1959).
- Flint, E. B. & Suslick, K. S. Sonoluminescence from nonaqueous liquids: emissions from small molecules. *J. Am. Chem. Soc.* **111**, 6987–6992 (1989).
- Suslick, K. S. & Flint, E. B. Sonoluminescence of alkali metal salts. *J. Phys. Chem.* **95**, 1484–1488 (1991).

- Suslick, K. S., Gawienowski, J. J., Schubert, P. F. & Wang, H. H. Alkane sonochemistry. *J. Phys. Chem.* **87**, 2299–2301 (1983).
- Suslick, K. S., Gawienowski, J. W., Schubert, P. F. & Wang, H. H. Sonochemistry in non-aqueous liquids. *Ultrasonics* **22**, 33–36 (1984).
- Yaws, C. L. *Handbook of Vapor Pressures* Vol. 3, 383 (Gulf, Houston, 1994).
- Hickling, R. Effects of thermal conduction in sonoluminescence. *J. Acoust. Soc. Am.* **35**, 967–974 (1963).
- Young, F. R. Sonoluminescence from water containing dissolved gases. *J. Acoust. Soc. Am.* **60**, 100–104 (1976).
- Yasui, K. *J. Phys. Soc. Jpn* **65**, 2830–2840 (1996).
- Suslick, K. S. & Flint, E. B. Sonoluminescence of non-aqueous liquids. *Nature* **330**, 553–555 (1987).
- Zel'dovich, Y. D. & Raizer, Y. P. *Physics of Shock Waves and High Temperature Hydrodynamic Phenomena* 2nd edn (Academic, New York, 1966).

Acknowledgements

This work was supported by the US National Science Foundation, the US Department of Energy, and in part by the US Defence Advanced Research Projects Agency.

Correspondence and requests for materials should be addressed to K.S.S. (e-mail: ksuslick@uiuc.edu).

Carbon cycling and chronology of climate warming during the Palaeocene/Eocene transition

Richard D Norris* & Ursula Röhl†

* MS-23, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02540-1541, USA

† Fachbereich Geowissenschaften, Universität Bremen, Postfach 330 440, 28334 Bremen, Germany

Current models of the global carbon cycle lack natural mechanisms to explain known large, transient shifts in past records of the stable carbon-isotope ratio ($\delta^{13}\text{C}$) of carbon reservoirs^{1,2}. The injection into the atmosphere of ~1,200–2,000 gigatons of carbon, as methane from the decomposition of sedimentary methane hydrates, has been proposed to explain a $\delta^{13}\text{C}$ anomaly^{3,4} associated with high-latitude warming¹ and changes in marine^{5–7} and terrestrial⁸ biota near the Palaeocene–Eocene boundary, about 55 million years ago. These events may thus be considered as a natural ‘experiment’ on the effects of transient greenhouse warming. Here we use physical, chemical and spectral analyses of a sediment core from the western North Atlantic Ocean to show that two-thirds of the carbon-isotope anomaly occurred within no more than a few thousand years, indicating that carbon was catastrophically released into the ocean and atmosphere. Both the $\delta^{13}\text{C}$ anomaly and biotic changes began between 54.93 and 54.98 million years ago, and are synchronous in oceans and on land. The longevity of the $\delta^{13}\text{C}$ anomaly suggests that the residence time of carbon in the Palaeocene global carbon cycle was ~120 thousand years, which is similar to the modelled response after a massive input of methane^{3,4}. Our results suggest that large natural perturbations to the global carbon cycle have occurred in the past—probably by abrupt failure of sedimentary carbon reservoirs—at rates that are similar to those induced today by human activity.

The carbon isotope record of marine carbonates is interrupted by a transient negative $\delta^{13}\text{C}$ anomaly of ~3‰ over an interval of ~100–200 kyr near the end of the Palaeocene^{1,9–13}. In the oceans, the $\delta^{13}\text{C}$ anomaly is associated with the extinction of 35–50% of cosmopolitan benthic foraminifera^{7,10}, the appearance of three short-ranging planktic foraminifera⁵, a widespread increase in

kaolinite (interpreted to herald a warm, wet climate), and a dramatic warming of high-latitude surface water and deep water (by $\sim 5\text{--}7^\circ\text{C}$) (ref. 1). The benthic foraminifer extinction and the rise in temperatures have been interpreted as indications of greenhouse warming and formation of oxygen-deficient deep waters¹. A $\delta^{13}\text{C}$ anomaly has also been observed in terrestrial soil carbonates and mammal teeth associated with a major turnover in land mammal assemblages^{8,14–16}. However, uncertainties in the number⁹ and correlation of the $\delta^{13}\text{C}$ events have led to estimates of their age ranging from 54.88 to 55.5 Myr (refs 8, 17–20). Here, we explore the age and duration of the $\delta^{13}\text{C}$ event using an astronomically calibrated timescale.

The Ocean Drilling Program (ODP) Site 1051 on Blake Nose in the western North Atlantic preserves an apparently unbroken record of magnetochron C24r and the events surrounding the Palaeocene–Eocene boundary²¹. In all, the record of chron C24r spans ~ 76 m of siliceous nannofossil chalk at ODP Site 1051, making this the thickest known sedimentary record of this magnetochron in the deep sea. We analysed records of iron (Fe) intensities (counts per second), measured by a non-destructive X-ray fluorescence (XRF) core scanner^{17,22}, and magnetic susceptibility (Fig. 1).

The benthic foraminifer extinction¹⁰ occurs at 512.80 m.c.d. (metres composite depth) in ODP Site 1051 and is closely associated with the first-appearance datum of the calcareous nannofossil *Rhombaster cuspis* that marks the base of biozone CP9a (ref. 21). Planktonic foraminifera characteristic of the $\delta^{13}\text{C}$ event (*Morozovella allisonensis*, *Acarinina sibiyaensis*, and *Acarinina africana*) appear over an interval, 3 m thick, immediately above the benthic extinction horizon (Fig. 2). These biostratigraphic events occur near the middle of planktic foraminifer biozone P5,

in agreement with other deep-sea sites^{5,20}, and verify that ODP Site 1051 is representative of the deep-sea record. Stable-isotope analyses of bulk carbonate from ODP Site 1051 demonstrate that the benthic foraminifer extinction and associated biostratigraphic events coincide with a $\sim 1.6\text{‰}$ $\delta^{13}\text{C}$ excursion (Fig. 2). Analyses of the planktic foraminifer *Morozovella subbotinae* show a 2.8‰ $\delta^{13}\text{C}$ decrease through the carbon isotope anomaly, which suggests that Site 1051 records the full magnitude of the excursion. The $\delta^{13}\text{C}$ anomaly in bulk carbonate records the sharp initial drop and more gradual recovery typical of correlative events in other deep sea sites, such as ODP Site 690 in the Southern Ocean^{1,2} and ODP Site 1001 in the Caribbean¹³.

Colour reflectance²¹, magnetic susceptibility, and Fe intensity data from ODP Site 1051 display well defined cycles throughout the upper Palaeocene (Fig. 1). Counts of Fe-intensity cycles reveal 23–25 cycles with chron C25n at ODP Site 1051. The uncertainty in the cycle count is due to a gap of 70 cm in discrete magnetic polarity measurements at the top of chron C25n. The estimated duration of chron C25n is ~ 487 kyr (ref. 20) which suggests that the Fe-intensity cycles are each 20–21 kyr long. The cycle duration is remarkably close to the modern mean precession period of ~ 21 kyr (ref. 23) that is the average of the 23- and 19-kyr precession bands.

Spectral analysis of the Fe intensity between 516 to 526 m.c.d. displays strong spectral power at wavelengths of 23 cm and 29 cm. The ratios of the Fe-intensity spectral peaks and their relative magnitudes are similar to those of the characteristic double peaks for the 19- and 23-kyr cycles in modern insolation spectra for the same latitude as Site 1051 (Fig. 3a and b). The Fe-intensity and magnetic-susceptibility cycles are probably due to variations in

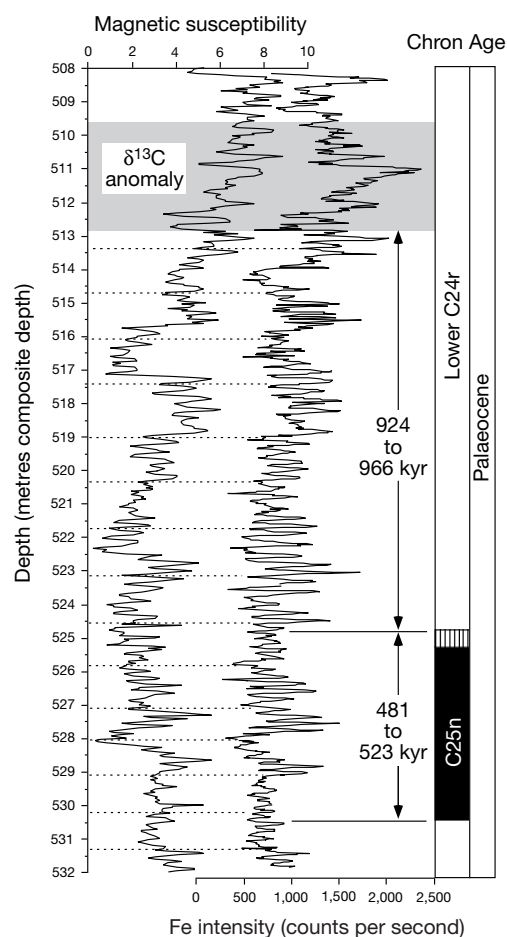


Figure 1 Cyclostratigraphy and magnetostratigraphy of the uppermost Palaeocene.

Magnetic susceptibility, Fe intensity and magnetostratigraphy from ODP Site 1051 on the Blake Nose ($30^\circ 03' \text{N}$; $75^\circ 21' \text{W}$, 1,983 m water depth). Ages are estimated from counts of physical property cycles and assumption of a duration of 21 kyr per cycle (see Fig. 3). Horizontal dashed lines bundle groups of five cycles. The depth scale is based on a spliced record of magnetic susceptibility and Fe intensity in holes 1051A and 1051B. The resulting composite record was cross-checked with downhole logging²¹ of magnetic susceptibility and FMS (formation microscanner) resistivity. The position of the $\delta^{13}\text{C}$ anomaly (Fig. 2) is shown as a shaded band. Post-cruise analysis of discrete samples demonstrate the existence of normal-polarity events in ODP Hole 1051A which have been confirmed as chrons C25n and C24n based on planktic foraminifera and calcareous nannofossil biostratigraphy as well as by correlation of downhole logs to the magnetic polarity record from nearby ODP Site 1050 (ref. 21). Uncertainty in the placement of the top of normal polarity chron C25n is shown by a striped band. The base of C25n is based upon shipboard pass-through magnetometer results²¹.

carbonate productivity and iron input from terrestrial sources, but may also reflect variations in carbonate dissolution.

Our cyclostratigraphy provides, to our knowledge, the first astronomically calibrated duration for the $\delta^{13}\text{C}$ anomaly. The primary-cycle wavelength increases from 23–29 cm below the $\delta^{13}\text{C}$ anomaly to ~51 cm within it, and increases again to ~105 cm above it (Fig. 3a). These power spectra maintain a double-peaked pattern above and below the $\delta^{13}\text{C}$ anomaly that we interpret as representing the main precession bands. The precession cycles seem to have been attenuated by a large increase in sedimentation rate above ~513 m.c.d. There appear to be a total of 7 magnetic susceptibility cycles between the onset of the anomaly and the point at which $\delta^{13}\text{C}$ reaches the post-anomaly average ratio (~509.8 m.c.d.; Fig. 2). The entire excursion thus lasted ~150 kyr.

Our estimate for the duration of the excursion has an error of ± 1.7 kyr, associated with the assumption of a 21-kyr average for seven precession cycles²³ and potentially larger errors associated with our identification of the start and end of the $\delta^{13}\text{C}$ anomaly. A debris flow of chalk clasts²¹ occurs just below the $\delta^{13}\text{C}$ anomaly and may have eroded several thousand years of sedimentation. In addition, we have chosen the termination of the $\delta^{13}\text{C}$ anomaly to be near its asymptote with post-excursion ratios, but a small shift in this level could change the duration of the anomaly by ~20 kyr or more.

Cycle counting demonstrates that the initial release of ^{12}C was very fast, and continued over an extended period. There is ~1 cycle between the base of the excursion and the point at which minimum $\delta^{13}\text{C}$ ratios are reached and a total of 1.5 cycles between the start of the event and the point at which $\delta^{13}\text{C}$ begins to rise again. The $\delta^{13}\text{C}$ of bulk carbonate drops abruptly from 2.2‰ to 0.96‰ at the start of the event and decreases more slowly over the next ~20 kyr before reaching near-minimum ratios of ~0.58‰. Minimum $\delta^{13}\text{C}$ levels

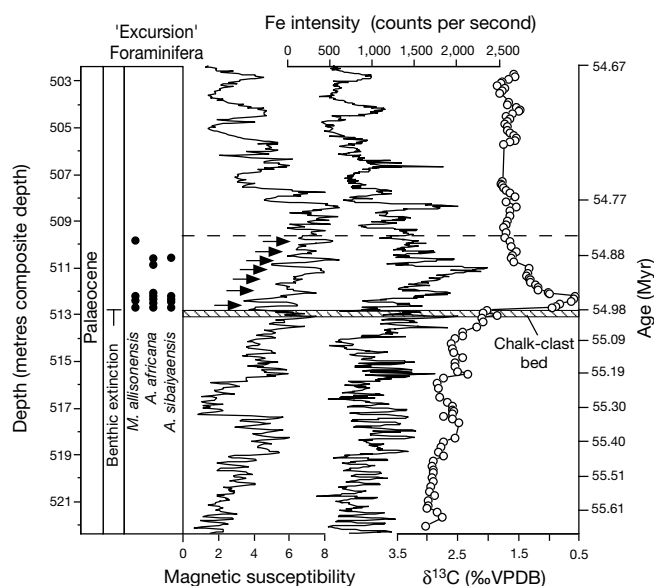


Figure 2 Stratigraphy around the $\delta^{13}\text{C}$ anomaly. Left to right: ranges of members of the 'excursion fauna' of planktic foraminifera (see text), magnetic susceptibility, Fe intensity, and $\delta^{13}\text{C}$ of bulk carbonate in ODP Site 1051. Small arrows show approximate positions of peaks in magnetic-susceptibility and Fe-intensity cycles believed to represent ~21-kyr precession cycles within the $\delta^{13}\text{C}$ anomaly. The dashed line shows the position of the termination of the $\delta^{13}\text{C}$ anomaly, and the hatched bar shows the position of the chalk-clast horizon. The age scale is based upon the locations of minima in magnetic susceptibility assuming an age of 54.98 Myr for the base of the $\delta^{13}\text{C}$ anomaly. Note the attenuation of the age scale due to an increase in sedimentation rate above the chalk-clast horizon. VPDB, Vienna Pee Dee Belemnite.

are maintained for as much as an additional 10 kyr. Our results are similar to isotopic measurements of benthic foraminifera and bulk carbonate from ODP Hole 690B that show a sharp drop in $\delta^{13}\text{C}$ followed by a more gradual shift to minimum $\delta^{13}\text{C}$ (refs 1, 2) levels.

The cyclostratigraphy suggests that about two-thirds of the $\delta^{13}\text{C}$ shift was completed catastrophically within a few thousand years or less. Given the large mass of carbon in the global carbon cycle (~42,000 Gton, ref. 3) it is hard to perturb the carbon-isotope composition of the oceans and atmosphere without a massive injection of isotopically light carbon, such as bacterial methane^{3,4}. Hence, our data strongly supports models for abrupt release of methane from submarine reservoirs by slope failure²⁻⁴. The extended period taken to reach minimum $\delta^{13}\text{C}$ suggests that methane, or another source of ^{12}C , bled into the oceans and atmosphere over ~30 kyr at a rate more than sufficient to balance its removal by burial. Gradual venting of ^{12}C -enriched material may reflect the time required to destabilize methane reservoirs in successively deeper waters through a feedback process. We suggest that the feedback system involved CO_2 -induced greenhouse warming of deep waters at their sites of formation followed by the collapse of additional hydrate reservoirs by hydrate melting and slope failure under warm deep water²⁻⁴. Any feedback between global warming and carbon release must have stopped about 30 kyr after it began, at

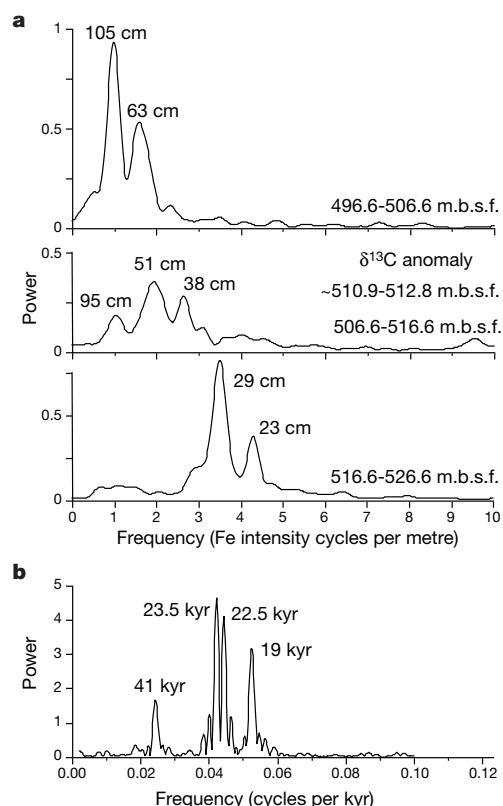


Figure 3 Power spectra for Site 1051 and modern summer insolation. **a**, Power spectra are shown for three 10-m windows in Site 1051. The double peak in spectra for 499.6–506.6 and 516.6–526.6 metres composite depth (m.c.d.) strongly suggests that the major cycles in the Fe-intensity and magnetic-susceptibility records reflect the 19- and 23-kyr precession bands. The increase in wavelength of the major Fe-intensity cycles suggests an increase in sedimentation rate within and above the $\delta^{13}\text{C}$ anomaly. The peak in spectral power at 95 cm in the 506.6–516.6 m.c.d. spectrum reflects the rapid shift from a dominant 51-cm cycle to an ~1-m cycle within this 10-m-thick interval. Power-depth spectra were created using the Blackman–Tukey method on 50-point detrended Fe-intensity records (Fig. 1). **b**, Power spectra for the modern summer insolation at 30° N. Modern insolation was calculated using the method of Berger²⁹. All spectra were generated with the AnalyseSeries software package³⁰. m.b.s.f., metres below sea floor.

which point $\delta^{13}\text{C}$ began its exponential rise toward the average post-excursion ratios.

There are 44–46 magnetic-susceptibility and Fe-intensity cycles of wavelength $\sim 23\text{--}29$ cm between the top of chron C25n and the beginning of the $\delta^{13}\text{C}$ anomaly at Site 1051 (Fig. 1). The cycle count suggests that the $\delta^{13}\text{C}$ anomaly began 924–966 kyr after the end of chron C25n. Cande and Kent²⁴ estimate an age of 55.9 Myr for the top of chron C25n, which places the start of the $\delta^{13}\text{C}$ anomaly at 54.93–54.98 Myr. These dates for the $\delta^{13}\text{C}$ anomaly are maximum ages, as some sediment may have been eroded during deposition of the chalk clast layer.

The 55.9 Myr data for the top of chron C25n may be significantly in error. The date is based on the calibration point of 55 Myr for the calcareous nannofossil zone NP9/NP10 boundary^{18,24} which, unfortunately, was established where the zonal boundary is an unconformity¹⁸. This 55-Myr date was used²⁹ to estimate the ages of the C24n and C25n polarity boundaries. As the NP9/N10 boundary is known to be younger than the $\delta^{13}\text{C}$ anomaly^{18,20}, the two events cannot have the same age and the 55.9-Myr age for the top of chron C25n must therefore be inaccurate. Accordingly, our absolute ages for events around the $\delta^{13}\text{C}$ anomaly will need to be revised when the ages of the magnetochrons are determined more precisely.

Our dates of $\sim 54.93\text{--}54.98$ Myr for the $\delta^{13}\text{C}$ anomaly and the associated climatic and biotic events show that these are synchronous with a $\delta^{13}\text{C}$ anomaly and biotic changes on land. A date of 54.96 Myr has been estimated for the prominent $\delta^{13}\text{C}$ excursion and the Wasatchian 0 provincial land-mammal age in the Bighorn basin, Wyoming, based on calculated sedimentation rates between the top of C25n (55.9 Myr) and the base of C24n (53.35 Myr) (ref. 25). Both the Bighorn-basin chronology and our chronology for Site 1051 are based on the same estimated ages²⁴ for the magnetic polarity events. We conclude that the age of the $\delta^{13}\text{C}$ anomaly relative to the top of chron C25n is easily within the same age range in the marine and terrestrial records, contrary to previous dates for the marine $\delta^{13}\text{C}$ event^{18,20}.

The synchronicity of the $\delta^{13}\text{C}$ event in the ocean and atmosphere unites the immigration of mammals such as artiodactyls, perissodactyls and primates into North America^{8,26,27} with the warming of subtropical and high-latitude oceans and the development of widespread disaerobic conditions in the deep sea^{7,10}. The initial release of ^{12}C -enriched carbon occurred over no more than a few thousand years and is precisely coincident with the benthic foraminifer extinction and the origination of the “excursion fauna” of planktic foraminifera.

The Site 1051 cyclostratigraphy shows that the entire $\delta^{13}\text{C}$ anomaly occurred over a span of $\sim 150 \pm 20$ kyr. The $\delta^{13}\text{C}$ anomaly developed over ~ 30 kyr, whereas the carbon released during the initial phases of the excursion was sequestered by burial over the remaining ~ 120 kyr. Therefore, the residence time of carbon in the global carbon cycle was ~ 120 kyr in the Palaeocene. Our estimate of the time to sequester carbon by burial will need to be confirmed at other sites; however, it agrees with the calculated modern carbon residence time of 140 kyr after injection of 1,200 Gton of methane into the pre-industrial carbon cycle^{3,4}. We suggest that the rate of carbon cycling was not very different from today during warm times such as the Palaeocene and early Eocene.

Current carbon-cycle models do not explicitly include exchange fluxes for methane hydrates and other deeply buried carbon sources other than by anthropogenic release⁴. Yet modern oceanic hydrate reservoirs contain more than 14,000 Gton of methane²⁸, which is large compared to the remaining total exchangeable carbon reservoir ($\sim 42,000$ Gton, ref. 3). The size of the Palaeocene–Eocene carbon-isotope excursion and the rapidity of its onset demonstrates that there must be a large reservoir of exchangeable carbon that can be released by natural processes on timescales similar to modern rates of anthropogenic carbon input. □

Received 3 May; accepted 14 September 1999.

- Kennett, J. P. & Stott, L. D. Abrupt deep-sea warming, paleoceanographic changes and benthic extinctions at the end of the Paleocene. *Nature* **353**, 225–229 (1991).
- Bains, S., Corfield, R. M. & Norris, R. D. Mechanisms of climate warming at the end of the Paleocene. *Science* **285**, 724–727 (1999).
- Dickens, G. R., O'Neill, R. R., Rea, D. K. & Owen, R. M. Dissociation of oceanic methane hydrate as a cause of the carbon isotope excursion at the end of the Paleocene. *Paleoceanography* **10**, 965–971 (1995).
- Dickens, G. R. Methane oxidation during the late Paleocene thermal maximum. *Bull. Soc. Géol. Fr.* (in the press).
- Kelly, D. C., Bralower, T. J., Zachos, J. C., Permolli Silva, I. & Thomas, E. Rapid diversification of planktonic foraminifera in the tropical Pacific (ODP Site 865) during the late Paleocene thermal maximum. *Geology* **24**, 423–426 (1996).
- Thomas, E. in *Global Catastrophes in Earth History: an Interdisciplinary Conference on Impacts, Volcanism, and Mass Mortality* (eds Sharpton, V. L. & Ward, P.) 481–495 (Geological Society of America, Boulder, Colorado, 1990).
- Thomas, E. in *Late Paleocene-Early Eocene Climatic and Biotic Events* (eds Aubry, M.-P., Lucas, S. & Berggren, W. A.) 214–243 (Columbia Univ. Press, New York, 1998).
- Koch, P. L., Zachos, J. C. & Gingerich, P. D. Correlation between isotope records in marine and continental carbon reservoirs near the Palaeocene/Eocene boundary. *Nature* **359**, 319–322 (1992).
- Stott, L. D., Kennett, J. P., Shackleton, N. J. & Corfield, R. M. The evolution of Antarctic surface waters during the Paleogene: inferences from stable isotopic composition of planktonic foraminifera. *Proc. ODP Sci. Res.* **113**, 849–863 (1990).
- Thomas, E. & Shackleton, N. J. in *Correlation of the Early Paleogene in Northwest Europe* (eds Knox, R. W., Corfield, R. M. & Dunay, R. E.) 401–441 (Geological Society, London, 1996).
- Pak, D., Miller, K. G. & Browning, J. Global significance of an isotopic record from the New Jersey Coastal Plain: Linkage between the shelf and deep sea in the late Paleocene to early Eocene. *Proc. ODP Sci. Res.* **150**, 305–315 (1997).
- Lu, G. & Keller, G. (1993). The Palaeocene-Eocene transition in the Antarctic Indian Ocean: Inference from planktonic foraminifera. *Mar. Micropaleontol.* **21**, 101–142 (1993).
- Bralower, T. J. et al. High resolution records of the late Paleocene thermal maximum and circum-Caribbean volcanism: is there a causal link? *Geology* **25**, 963–966 (1997).
- Koch, P. L., Zachos, J. C. & Dettman, D. L. Stable isotope stratigraphy and paleoclimatology of the Paleogene Bighorn Basin (Wyoming, USA). *Palaeogeogr., Palaeoclimatol., Palaeoecol.* **115**, 61–89 (1995).
- Clyde, W. C. & Gingerich, P. D. Mammalian community response to the latest Palaeocene thermal maximum: An isotaphonomic study in the northern Bighorn Basin, Wyoming. *Geology* **26**, 1011–1014 (1998).
- Hooker, J. J. in *Late Paleocene-Early Eocene Climatic and Biotic Events* (eds Aubry, M.-P., Lucas, S. & Berggren, W. A.) 428–450 (Columbia Univ. Press, New York, 1998).
- Röhl, U. & Abrams, L. J. in *Proc. ODP Sci. Res.* (eds Leckie, R. M. et al.) Vol. 165 (Ocean Drilling Program, College Station, Texas, in the press).
- Aubry, M.-P., Berggren, W. A., Stott, L. D. & Sinha, A. in *Correlation of the Early Paleogene in Northwest Europe* (eds Knox, R. W., Corfield, R. M. & Dunay, R. E.) 353–380 (Geological Society, London, 1996).
- Wing, S. L. in *Late Paleocene-Early Eocene Climatic and Biotic Events* (eds Aubry, M.-P., Lucas, S. & Berggren, W. A.) 380–400 (Columbia Univ. Press, New York, 1998).
- Berggren, W. A., Kent, D. V., Swisher, C. C. III & Aubry, M.-P. in *Geochronology, Time Scales and Global Stratigraphic Correlations: A Unified Temporal Framework for an Historical Geology* (eds Berggren, W. A., Kent, D. V., Aubry, M.-P. & Hardenbol, J.) 129–212 (Society for Economic Paleontologists and Mineralogists, Tulsa, Oklahoma, 1995).
- Norris, R. D. et al. Black Nose Paleoceneanographic Transect, Western North Atlantic. *Proc. ODP Init. Rep. B* **171**, 1–749 (1998).
- Jansen, J. H. F., Van der Gaast, S. J., Koster, B. & Vaars, A. J. CORTEX, a shipboard XRF-scanner for element analyses in split sediment cores. *Mar. Geol.* **151**, 143–153 (1998).
- Herbert, T. D., Permolli Silva, I., Erba, E. & Fischer, A. G. in *Geochronology, Time Scales and Global Stratigraphic Correlation* (eds Berggren, W. A., Kent, D. V., Aubry, M.-P. & Hardenbol, J.) 81–93 (Society for Economic Paleontologists and Mineralogists, Tulsa, Oklahoma, 1995).
- Cande, S. C. & Kent, D. V. Revised calibration of the geomagnetic polarity time scale for the Late Cretaceous and Cenozoic. *J. Geophys. Res.* **100**, 6093–6095 (1995).
- Wing, S. L., Boer, H. & Koch, P. L. in *Warm Climates in Earth History* (eds Huber, B. T., MacLeod, K. & Wong, S. L.) (Cambridge Univ. Press, Cambridge, UK, in the press).
- Wing, S. L., Alroy, J. & Hickey, L. J. Plant and mammal diversity in the Paleocene to early Eocene of the Bighorn Basin. *Palaeogeogr., Palaeoclimatol., Palaeoecol.* **115**, 117–155 (1995).
- Maas, M., Anthony, M. R. L., Gingerich, P. D., Gunnell, G. F. & Krause, D. W. Mammalian genetic diversity and turnover in the Late Paleocene and early Eocene of the Bighorn and Crazy Mountains Basins, Wyoming and Montana (USA). *Palaeogeogr., Palaeoclimatol., Palaeoecol.* **115**, 181–207 (1995).
- Gornitz, V. & Fung, I. Potential distribution of methane hydrates in the world's oceans. *Glob. Biogeochem. Cycles* **8**, 335–347 (1994).
- Berger, A. Long-term variations of daily insolation and quaternary climatic change. *J. Atmos. Sci.* **35**, 2362–2367 (1978).
- Paillard, D., Labeyrie, L. & Yiou, P. Macintosh program performs time-series analysis. *Eos* **77**, 379 (1996).

Acknowledgements

We thank G. Dickens, M. P. Aubry, W. A. Berggren, D. Kent, P. Koch and D. Kroon for discussions, and the staff of the Ocean Drilling Program in Bremen for technical assistance. This work was supported by the Joint Oceanographic Institutions—US Science Support Advisory Committee (JOI-USSAC), and the National Science Foundation and the Deutsche Forschungsgemeinschaft.

Correspondence and requests for materials should be addressed to R.D.N. (e-mail: rnmorris@whoi.edu).

Onset of permanent stratification in the subarctic Pacific Ocean

Gerald H. Haug^{*†}, Daniel M. Sigman[†], Ralf Tiedemann[‡], Thomas F. Pedersen[§] & Michael Sarnthein^{||}

^{*} Department of Earth Sciences, University of Southern California, Los Angeles, California 90089-0740, USA

[†] Department of Geosciences, Princeton University, Princeton, New Jersey 08544, USA

[‡] GEOMAR, 24148 Kiel, Germany

[§] Department of Earth and Ocean Sciences, University of British Columbia, Vancouver, V6T 1Z4 Canada

^{||} Institut für Geowissenschaften, Universität Kiel, 24098 Kiel, Germany

[¶] Present address: Geological Institute, ETH, 8092 Zürich, Switzerland.

The surface waters of the modern subarctic Pacific Ocean are isolated from the nutrient-rich waters below by a steep vertical gradient in salinity (halocline), a feature which is a dominant control on upper-ocean stratification in polar environments^{1–3}. The physical processes which maintain the halocline and, in turn, its physical, biological, and geochemical effects have long been subjects of intense inquiry^{3,4}. The stratification of polar surface waters influences the exchange of CO₂ between ocean and atmosphere^{5–8}, so the history of the subarctic Pacific halocline may have played a role in past changes in atmospheric CO₂ concentration. Here we report opal accumulation rates and nitrogen-isotope data from sediments in this region which indicate that the subarctic Pacific halocline developed abruptly 2.73 million years ago, coincident with the onset of extensive Northern Hemisphere glaciation. The halocline would have reduced the transport of nutrient-rich deep water into the euphotic zone, leading to a decrease in biological production but an increase in the fraction of nutrient stocks utilized. This increase in the efficiency of the 'biological pump' would have lowered the rate of CO₂ evasion from ocean to atmosphere, potentially reducing atmospheric CO₂ concentrations from the suggested higher level of the preceding mid-Pliocene warm interval^{9,10}.

In the large low- to mid-latitude gyres of the world ocean, the major nutrients, nitrate, phosphate and silicate, are completely consumed by biological production. However, in the Southern Ocean and the subarctic Pacific, non-zero nutrient concentrations occur in surface waters. The cause of incomplete nutrient consumption in these high-nutrient regions is still uncertain, with limitations on phytoplankton growth rate playing an important role⁴. However, one fundamental characteristic of these regions is a large flux of nutrient-rich deep water into the euphotic zone due to Ekman divergence and vertical mixing¹¹. In the Southern Ocean today, high rates of deep water exposure at the surface overwhelm the capacity of phytoplankton to consume nutrients in the course of growth. As a result, nitrate utilization (the ratio of nitrate uptake to nitrate supply) is only ~30% in the Antarctic zone^{8,12}.

The subarctic Pacific, although it is also a high-nutrient polar regime, maintains a much higher degree of nutrient utilization than the Southern Ocean. In the open subarctic Pacific, summer NO₃⁻ concentrations frequently fall below 8 µM in surface water, despite nitrate concentrations of greater than 35 µM in the upwelling deep water below the permanent halocline¹³. One important difference between the subarctic Pacific and the Antarctic Ocean is the much stronger halocline in the former^{1–3}, where the extremely low salinity

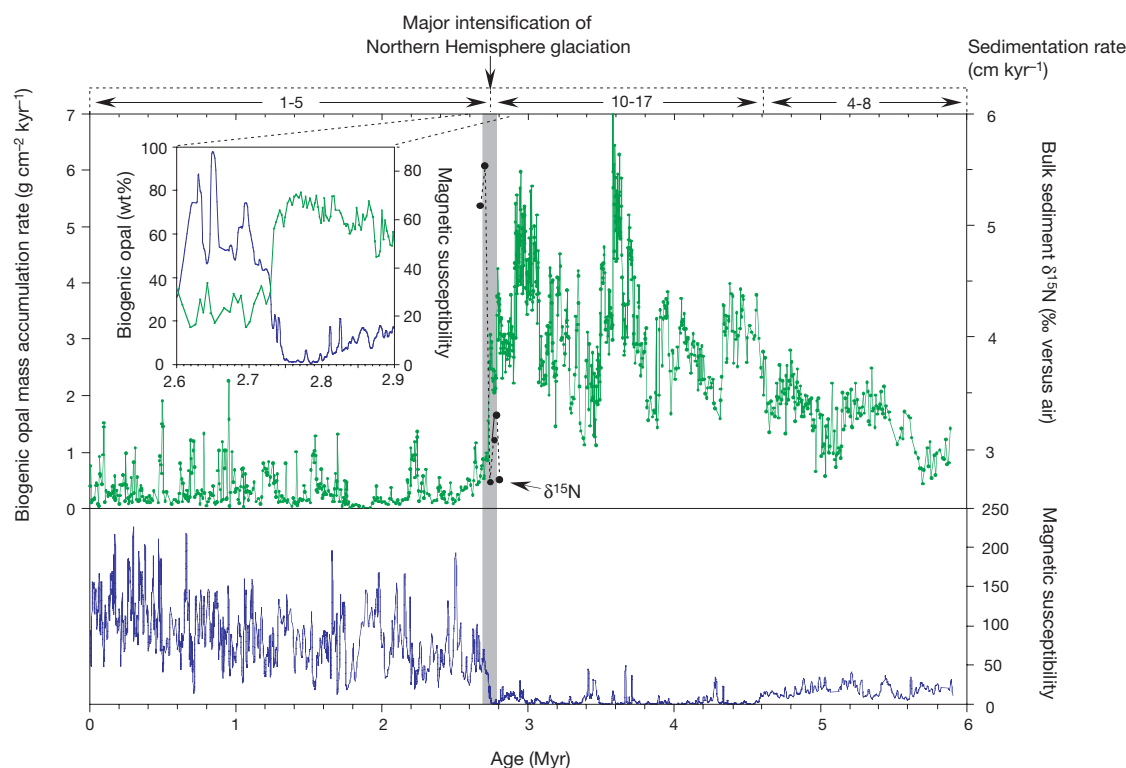


Figure 1 Subarctic Pacific palaeoceanographic time series (ODP Site 882). Bottom panel, fluctuations in magnetic susceptibility (in c.g.s., data smoothed by a 5-point running mean) reflect the input of ice-rafted debris (blue line)¹⁴. Middle panel, mass accumulation rates of biogenic opal (green line) and bulk sediment $\delta^{15}\text{N}$ (black data points). Top panel,

changes in sedimentation rates. Inset, fluctuations in magnetic susceptibility (blue line) and biogenic opal (green line) during the time interval 2.6–2.9 Myr. $\delta^{15}\text{N}(\text{‰}) = ((^{15}\text{N}/^{14}\text{N})_{\text{sample}} / (^{15}\text{N}/^{14}\text{N})_{\text{standard}} - 1) \times 1,000$, where N₂ in air is the standard.

of the upper 400 m limits the exposure of nutrient- and CO₂-charged subsurface waters at the surface.

The production of biogenic opal is closely related to the availability of nutrients, with high opal production being limited to regions where silicate-rich deep water is supplied to diatoms growing in sunlit surface waters¹⁴. In order to reconstruct changes in the nutrient status of the subarctic Pacific, we have generated a 6-Myr biogenic opal record from Ocean Drilling Program (ODP) Site 882 (50° 21' N, 167° 35' E; water depth 3,244 m), determined by the density separation technique¹⁵. The timescale at Site 882 is based on palaeomagnetic reversals for initial age control¹⁶ and an astronomically calibrated age model between the magnetic reversals¹⁷.

The most pronounced feature of the Site 882 opal record is an approximately fourfold decrease in opal accumulation rate at 2.73 Myr, from 3–6 g cm⁻² kyr⁻¹ to <1.5 g cm⁻² kyr⁻¹. This abrupt drop in opal accumulation occurred at isotope stage G6 (the former isotope stage 110), at the same time as the appearance of abundant ice-rafted debris (Fig. 1). Observations from sediment cores from the Gulf of Alaska¹⁶ indicate that this change occurred across the entire subarctic Pacific. In our 6-Myr record, opal accumulation was highest between 2.73 and 4.6 Myr (2–6 g cm⁻² kyr⁻¹), with intermediate rates between 4.6 and 6 Myr (0.7–2.5 g cm⁻² kyr⁻¹) (see Fig. 2). During the past 2.73 Myr, opal accumulation has been <2 g cm⁻² kyr⁻¹ (generally 0.1–1.2 g cm⁻² kyr⁻¹). Maxima in opal accumulation during the past 2.73 Myr are generally linked to interglacial times or deglaciations, as reported for the last deglaciation¹⁸.

Because silicate is supplied to the surface from nutrient-rich deep water, the high opal flux of the mid-Pliocene subarctic Pacific requires a rate of deep water exposure at the surface of similar magnitude to that observed in the modern Antarctic. Because silicate consumption is nearly complete in the modern subarctic Pacific, the sharp drop in opal flux at 2.73 Myr cannot be attributed simply to a decrease in the completeness of silicate consumption. Instead, it must record a decrease in the rate of exposure of silicate-rich deep water at the subarctic Pacific surface.

In the modern ocean, silicate consumption is complete over a

broad range of deep water upwelling rates. Thus, even with very high rates of deep water supply in the Pliocene subarctic Pacific, silicate utilization may well have been complete. However, nitrate and phosphate consumption is often incomplete in regions of high nutrient supply. For instance, in the modern Antarctic, where deep-water exposure rates are high, silicate is almost completely consumed by diatoms as surface water flows northward, but the degree of nitrate and phosphate utilization does not exceed 30% (ref. 8). By analogy, we would expect that the degree of nitrate and phosphate utilization was low in the mid-Pliocene subarctic Pacific and increased with the reduction in deep water exposure at 2.73 Myr.

The nitrogen isotopic composition of the sedimentary organic matter is used as a palaeoceanographic proxy for the degree of nitrate utilization, that is, the fraction of the nitrate supply which is consumed by phytoplankton^{19,20}. Phytoplankton consume ¹⁴NO₃⁻ preferentially over ¹⁵NO₃⁻; resulting in lower sedimentary δ¹⁵N values under nutrient-replete conditions. At Site 882, sediment δ¹⁵N values increase concurrently with the sharp drop in opal accumulation at 2.73 Myr (from ~3‰ to 5‰ versus air N₂; see Fig. 1). This shift towards higher δ¹⁵N values suggests an increase in nitrate utilization, confirming that the decrease in opal flux resulted from a decrease in the exposure of nutrient-rich deep water at the surface, which lowered the nitrate supply and thus forced more complete nitrate consumption while decreasing the silicate available for opal production. In contrast, a decrease in production without a decrease in nitrate supply would have resulted in less complete nitrate consumption and thus lower sediment δ¹⁵N values. The amplitude of the δ¹⁵N change at 2.73 Myr is large relative to measured glacial–interglacial variations in Pleistocene records from the western subarctic Pacific²¹, confirming that the 2.73-Myr event represented a fundamental and permanent shift in the nutrient status of the subarctic Pacific.

The Late Cenozoic closure of the Panamanian Seaway caused profound changes in deep Atlantic circulation at 4.6 Myr (refs 22, 23). Since 4.6 Myr, carbonate dissolution and epibenthic δ¹³C records in the deep equatorial west Atlantic Ocean (ODP sites

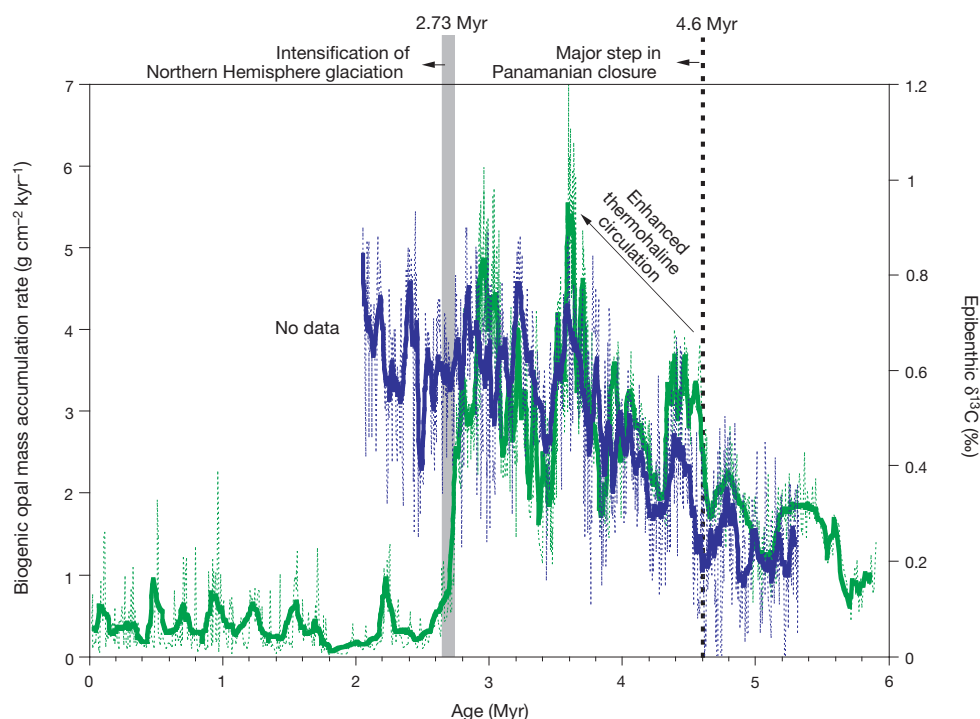


Figure 2 Subarctic Pacific and Caribbean Sea palaeoceanographic records. Shown is a comparison of biogenic opal mass accumulation rates (green line) at ODP Site 882 in the subarctic northwest Pacific and epibenthic δ¹³C values of *C. wuellerstorfi* (blue line) at

ODP Site 999 in the Caribbean sea¹⁹, plotted versus age. Thick lines are records smoothed by a 15-point running mean. δ¹³C(‰) = ((¹³C/¹²C)_{sample} / (¹³C/¹²C)_{standard} - 1) × 1,000; standard, PDB.

925–927, Ceara rise depth transect²²) and in the Caribbean Sea (ODP Site 999; ref. 23) (Fig. 2) record enhanced deep- and intermediate-water formation in the North Atlantic. The epibenthic (*Cibicoides wuellerstorfi*) $\delta^{13}\text{C}$ record from Site 999 shows a remarkable correlation with the Site 882 opal record between 5.3 and 2.73 Myr (Fig. 2). The general trend towards higher $\delta^{13}\text{C}$ values at Site 999 (increased ventilation in the Atlantic–Caribbean) coincides with the increase in biogenic opal at Site 882, and this correlation is also observed in the smaller, more frequent variations. This suggests that the strength of Atlantic-sourced deep circulation modulated the nutrient supply to the subarctic Pacific surface between 5.3 and 2.73 Myr.

The nutrient concentration of Pacific deep water is controlled on the large scale by the routes of deep circulation. Today, deep water originates in the North Atlantic and the Southern Ocean, and becomes progressively enriched in nutrients as it flows northward in the Pacific¹⁴. Enhanced deep-ocean circulation of this form appears to have gradually strengthened the interbasin gradient in deep-ocean nutrients since 4.6 Myr, lowering nutrient concentrations in the Atlantic and raising them in the North Pacific²⁴. Because of vertical exchange and upwelling in the subarctic Pacific, the strengthened interbasin gradient increased the supply rate of silicate to the subarctic Pacific surface, leading to high opal accumulation rates in the underlying sediments. In addition, upwelling in the subarctic Pacific represents a potential route by which deep water is converted back into surface- and intermediate-depth waters to complete the global thermohaline conveyor¹⁴. While the modern

transport of deep water into the surface layer is limited by the subarctic Pacific halocline, upwelling into the high-latitude North-Pacific surface may have been an important fate for Pacific deep and intermediate waters during the Pliocene.

However, deep circulation, as recorded by epibenthic $\delta^{13}\text{C}$ at Site 999 (ref. 23) and other locations²⁵, shows no abrupt change at 2.73 Myr, when opal accumulation decreased sharply at Site 882 (Fig. 2). Thus, there is no evidence for a radical change in the Atlantic–Pacific gradient in nutrient concentrations at this time. In addition, the silicate concentration in the deep North Pacific, which holds the oldest water in the ocean interior, is currently the highest observed throughout the ocean. Thus, the observed decrease in opal production at 2.73 Myr cannot readily be explained by a decrease in the circulation-driven enrichment of nutrients in the deep North Pacific. Moreover, aeolian inputs give no sign of a decrease in wind strength at 2.73 Myr (ref. 16), so the abrupt decrease in nutrient supply at that time cannot be attributed to decrease in Ekman-driven upwelling.

We propose that the sudden decrease in silicate and nitrate supply to the subarctic Pacific surface at 2.73 Myr resulted from the abrupt development of the strong subarctic Pacific halocline that dominates the region today. Before 2.73 Myr, the preglacial subarctic Pacific was perhaps analogous to the modern Antarctic zone of the Southern Ocean, where winds drive intense upwelling and vertical mixing, leading to high opal production rates and high surface nitrate and phosphate concentrations. With the onset of extensive glaciation, indicated by the increased supply of ice-rafted debris into the North Pacific, the development of a more intense permanent halocline caused the nutrient regime of the subarctic Pacific to diverge from that of the Antarctic, towards its modern conditions of reduced deep water exposure and lower surface-nutrient concentrations. There is precedent for this hypothesis in the inference of ref. 26, based on diatom assemblage data, that the upper water column of the far North Pacific was less stratified and more saline before Northern Hemisphere glaciation, although the timing of the evolution to modern conditions was uncertain.

The modern state of the North Pacific appears to be self-sustaining, with low surface salinity preventing the large-scale production of new subsurface waters; there is thus no need for a lateral input of surface water from the salty subtropical gyre to the south, allowing for the continued oceanographic isolation of the subarctic Pacific surface³. We do not know what specific change in the atmosphere or surface ocean triggered the sudden development of the subarctic Pacific halocline. However, this change was probably an integral part of the onset of significant Northern Hemisphere glaciation. The modern subarctic Pacific surface owes its low salinity to an excess in precipitation relative to evaporation, runoff from the continents, and the seasonal sea-ice cycle³. The low temperature of the region is an important factor driving of these inputs, and the increase in ice-rafted debris and $\delta^{18}\text{O}$ of planktonic foraminifera²⁷ at Site 882 do indeed argue for a large cooling of the region at, or just before, the onset of Northern Hemisphere glaciation at 2.73 Myr.

The exposure of nutrient- and CO_2 -rich deep water at the ocean surface releases CO_2 to the atmosphere. Conversely, nutrient and inorganic-carbon uptake by phytoplankton in surface waters, and export of the resultant organic matter, draws CO_2 out of the atmosphere and surface ocean, sequestering this carbon in the deep sea. Thus, nutrient utilization, the ratio of nutrient uptake to nutrient supply, corresponds to a similar ratio of CO_2 uptake to CO_2 release by the surface ocean. As a result, it describes the efficiency with which the biological pump removes CO_2 from the atmosphere: a higher degree of nutrient utilization lowers atmospheric CO_2 (refs 5–7). The $\delta^{15}\text{N}$ data from the subarctic Pacific suggest a low degree of nitrate utilization before 2.73 Myr. If this was so, the mid-Pliocene subarctic Pacific would have represented a significant leak in the biological pump, releasing deep-sequestered CO_2 back into the atmosphere. The subarctic Pacific

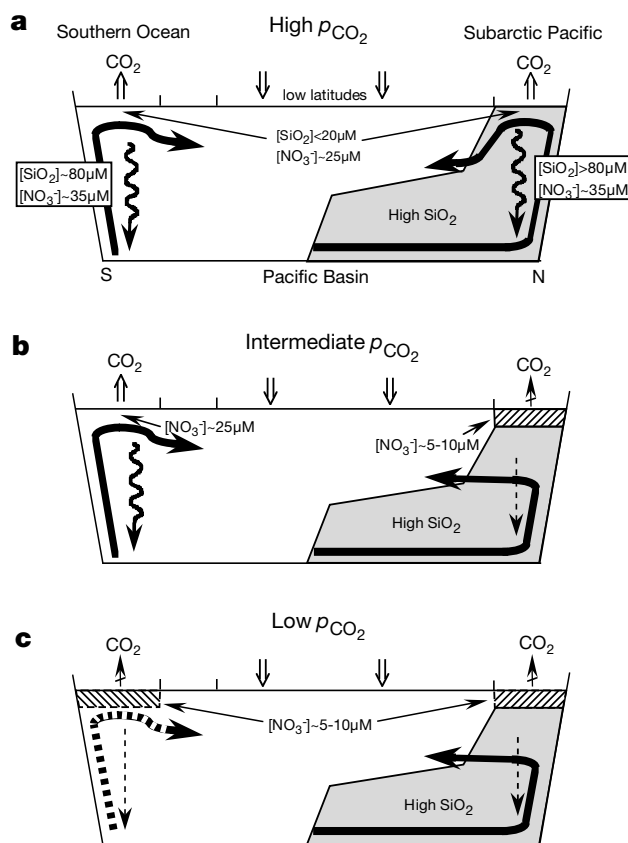


Figure 3 Schematic changes in the efficiency of the biological pump. Here we show a cross-section through the Pacific Ocean to illustrate the effects of changes in deep-water exposure and nutrient utilization on atmospheric CO_2 . **a**, Preglacial >2.73 Myr, bipolar deep water exposure; **b**, Pleistocene interglacial periods, unipolar Southern Ocean deep water exposure; **c**, Pleistocene glacial periods, stratified polar oceans; see text. Changes in temperature-driven solubility and CaCO_3 production and their effect on atmospheric CO_2 are not included here, as both are uncertain.

may thus have been important in the hypothesized greenhouse forcing of the warm mid-Pliocene climate.

The data from the subarctic Pacific are intriguing in the light of palaeoceanographic evidence for a Pleistocene oscillation in stratification and nutrient utilization in the Southern Ocean^{8,28,29}. During the last glacial period, the rate of deep water exposure was lower in the Antarctic zone⁸, perhaps because of the development of a stronger halocline³⁰. Despite less biological production, this decrease in the entrainment of nitrate- and CO₂-rich deep water apparently led to a net increase in nitrate utilization^{8,12,31}. The consequent reduction of CO₂ evasion from the Southern Ocean to the atmosphere would have lowered glacial atmospheric CO₂ concentrations^{5–8}.

Just as atmospheric CO₂ would have decreased during Pleistocene glacial periods with the stratification of the Antarctic surface ocean, the development of the subarctic Pacific halocline at 2.73 Myr would have reduced atmospheric concentrations of CO₂. This suggests the central importance of vertical stability in the high-latitude regions as a driver of atmospheric CO₂ changes (Fig. 3). Before 2.73 Myr, deep water exposure was apparently intense in both the subarctic Pacific and the Southern Ocean, leading to incomplete utilization of nutrients and CO₂ evasion in both polar regions (Fig. 3a). During the Pleistocene, there appears to have been two modes of polar vertical stability. During glacial times, both the subarctic Pacific and the Antarctic were vertically stable, preventing CO₂ loss to the atmosphere and resulting in low atmospheric CO₂ concentrations (Fig. 3c). During interglacial times (such as today), the subarctic Pacific was “capped” by its permanent halocline, but the upwelling rate was high in the Southern Ocean, leading to intermediate levels of atmospheric CO₂ (Fig. 3b).

Carbon-cycle model experiments for the Southern Ocean³¹ provide a basis for rough estimation of the effect of subarctic Pacific halocline development on Pliocene atmospheric CO₂ levels. If we consider the area and overturning rate of the Pliocene subarctic Pacific to have been roughly half that of the Antarctic, then the ~75% reduction in overturning suggested by the opal accumulation change would have caused a decrease in atmospheric CO₂ concentration of ~30–40 p.p.m., considering the effect of the biological pump alone and making a number of simplifying assumptions³¹. Quantification of this effect, and a physical explanation for the abrupt development and subsequent persistence of the subarctic Pacific halocline, remain to be achieved. □

Received 21 January; accepted 26 August 1999.

- Reid, J. L. Sea surface temperature, salinity, and density of the Pacific Ocean in summer and in winter. *Deep-Sea Res.* **16** (suppl.), 215–224 (1969).
- Talley, L. D. Distribution and formation of North Pacific intermediate water. *J. Phys. Oceanogr.* **23**, 517–537 (1993).
- Warren, B. Why is no deepwater formed in the North Pacific? *J. Mar. Res.* **41**, 327–347 (1983).
- Chisholm, S. W. & Morel, F. M. M. (eds) What controls phytoplankton production in nutrient-rich areas in the open sea? *Limnol. Oceanogr.* **36**, 1507–1970 (1991).
- Knox, F. & McElroy, M. B. Changes in atmospheric CO₂: Influence of the marine biota at high latitude. *J. Geophys. Res.* **89**, 4629–4637 (1984).
- Sarmiento, J. L. & Toggweiler, R. A new model for the role of the oceans in determining atmospheric pCO₂. *Nature* **308**, 621–624 (1984).
- Siegenthaler, U. & Wenk, T. Rapid atmospheric CO₂ variations and ocean circulation. *Nature* **308**, 624–626 (1984).
- Francois, R. et al. Contributions of Southern Ocean surface water stratification to low atmospheric CO₂ concentrations during the last glacial period. *Nature* **389**, 929–935 (1997).
- Budyko, M. I., Ronov, A. B. & Yanshin, A. I. The history of the Earth atmosphere. *Gidrometeoizdat* 1–139 (Springer, Berlin; English translation, 1987).
- Raymo, M. E., Grant, B., Horowitz, M. & Rau, G. H. Mid-Pliocene warmth: stronger greenhouse and stronger conveyor. *Mar. Micropaleontol.* **27**, 313–326 (1996).
- Gargett, A. E. Physical processes and the maintenance of nutrient-rich euphotic zones. *Limnol. Oceanogr.* **36**, 1527–1546 (1991).
- Falkowski, P. G., Barber, R. T. & Smetacek, V. Biogeochemical controls and feedbacks on ocean primary production. *Science* **281**, 200–206 (1998).
- Levitus, S., Conkright, M., Reid, J. L., Najjar, R. G. & Mantyla, A. Distribution of nitrate, phosphate and silicate in the world ocean. *Proc. Oceanogr.* **31**, 245–273 (1993).
- Broecker, W. S. & Peng, T. H. *Tracers in the Sea* 1–690 (Eldigio, New York, 1982).
- Haug, G. H. Zur Paläoceanographie und Sedimentationsgeschichte des Nord-Pazifiks während der letzten 6 Millionen Jahre (ODP Site 882, Leg 145). *Berichte-Reports Geol. Inst. Univ. Kiel* **78**, 1–98 (1996).

- Rea, D. K. et al. *Proc. ODP Int. Rep.* **145**, 85–119 (1993).
- Tiedemann, R. & Haug, G. H. Astronomical calibration of Site 882 cycle stratigraphy in the northwest. *Proc. ODP Sci. Res.* **145**, 283–293 (1995).
- Keigwin, L. D., Jones, G. A. & Froelich, P. N. A 15,000 year paleoenvironmental record from the Meiji Seamount, far Northwestern Pacific. *Earth Planet. Sci. Lett.* **111**, 425–440 (1992).
- Altabet, M. & Francois, R. Sedimentary nitrogen isotopic ratio as a recorder for surface ocean nitrate utilization. *Glob. Biogeochem. Cycles* **8**, 103–116 (1994).
- Farrell, J. W., Pedersen, T. F., Calvert, S. E. & Nielsen, B. Glacial–interglacial changes in nutrient utilization in the equatorial Pacific Ocean. *Nature* **377**, 514–517 (1995).
- Sigman, D. M., Keigwin, L. D., Altabet, M. A. & McCorkle, D. C. The nitrogen isotopic analysis of a deglacial productivity event in the western subarctic Pacific. *Eos* **74**, 185 (1993).
- Tiedemann, R. & Franz, S. O. Deep-water circulation, chemistry, and terrigenous sediment supply in the equatorial Atlantic during the Pliocene, 3.3–2.6 Ma and 5–4.5 Ma. *Proc. ODP Sci. Res.* **154**, 299–318 (1997).
- Haug, G. H. & Tiedemann, R. Effect of the formation of the Isthmus of Panama on Atlantic Ocean thermohaline circulation. *Nature* **393**, 673–676 (1998).
- Keir, R. On the Late Pleistocene ocean geochemistry and circulation. *Paleoceanography* **3**, 413–447 (1988).
- Raymo, M. E., Hodell, D. & Jansen, E. Response of deep ocean circulation to initiation of Northern Hemisphere Glaciation (3–2 Ma). *Paleoceanography* **7**, 645–672 (1992).
- Sancetta, C. & Silvestri, S. Pliocene–Pleistocene evolution of the North Pacific ocean–atmosphere system, interpreted from fossil diatoms. *Paleoceanography* **1**, 163–180 (1986).
- Maslin, M. A., Haug, G. H., Sarnthein, M., Tiedemann, R., Erlenkeuser, H. & Stax, R. Northwest Pacific Site 882: The initiation of northern hemisphere glaciation. *Proc. ODP Sci. Res.* **145**, 315–333 (1995).
- Mortlock, R. A. et al. Evidence for lower productivity in the Antarctic Ocean during the last glaciation. *Nature* **351**, 220–223 (1991).
- Kumar, N. et al. Increased biological productivity and export production in the glacial Southern Ocean. *Nature* **378**, 675–680 (1995).
- Morley, J. J. & Hayes, J. D. Oceanographic conditions associated with high abundances of the radiolarian *Cycladophora davisiana*. *Earth Planet. Sci. Lett.* **66**, 63–72 (1983).
- Sigman, D. M. et al. The isotopic composition of diatom-bound nitrogen in Southern Ocean sediments. *Paleoceanography* **14**, 118–134 (1999).

Acknowledgements

We thank J. Barron, R. Francois, S. Honjo, T. Joyce, L. Keigwin, M. Maslin and R. Toggweiler for discussions. This work was supported by the Deutsche Forschungsgemeinschaft (DFG) and the Natural Sciences and Engineering Research Council of Canada.

Correspondence and request for materials should be addressed to G.H.H. (e-mail: haug@erdw.ethz.ch).

The formation of Mount Etna as the consequence of slab rollback

Zohar Gvirtzman*† & Amos Nur*

* Geophysics Department, Stanford University, Stanford, California 94305, USA

Mount Etna, the largest volcano in Europe, lies close to the subduction-related Aeolian magmatic arc but shows no trace of subducted material in its magmas. Mount Etna is also situated on continental crust yet shows oceanic basalt affinities^{1–3}, with isotopic ratios of helium and carbon suggesting that it is fed by the same type of mantle source as are mid-ocean ridge basalts^{4,5}. Here we propose that although this giant volcano is not subduction-related—in the sense that it is not part of the magmatic arc—its formation is strongly related to the nearby subduction process. Based on a three-dimensional model of the tectonic plates in this region, we propose that the voluminous melting under Mount Etna results from ‘suction’ of asthenospheric material from under the neighbouring African plate. Such lateral flow is expected when descending slabs migrate backwards in the mantle (rollback) leaving low-pressure regions behind^{6,7} them. This was previously identified at the northern end of the Tonga arc (southwest Pacific Ocean) where such flow feeds arc⁸ or backarc⁹ magmatism. Here

† Present address: Institute of Earth Sciences, Hebrew University, Givat Ram, Jerusalem 91904, Israel.

we show that in the south Tyrrhenian subduction zone, slab rollback pulls asthenospheric material much farther along the plate contact, reaching the base of the crust in the forearc region. This explains the voluminous melting under Mount Etna and also the recent uplift of the forearc region (the Calabrian peninsula)¹⁰.

Figure 1 shows that Mount Etna is situated a little to the side of the Ionian slab, rather than directly above it. Also, it does not face the magmatic arc (Aeolian islands) but is situated a little forward, in line with the forearc region (Calabria), where the top of the slab, at ~70 km depth, is too shallow to produce melting. This location is consistent with the geochemical data suggesting that Mount Etna is not fed by material coming from the mantle wedge between the subducting and overriding plates. On the other hand, if Mount Etna is not related to the south Tyrrhenian subduction zone, why did it form so close to this region and approximately at the same time? Previous studies¹¹ show that Mount Etna is located at a junction between three fault systems, but this explains only how magma rises through the crust, not the cause for melting. Here we explain the relationship between Mount Etna and the nearby subduction zone, and particularly, suggest the cause for the voluminous melting at this specific location.

Our suggestion that melting under Mount Etna is caused by 'suction' of asthenospheric material from under the African plate is based on a three-dimensional model which we developed for the south Tyrrhenian subduction zone (shown schematically in Fig. 2 and to true scale in Fig. 3). This model describes the shape of the asthenospheric wedge between the subducting Ionian lithosphere and the overriding Tyrrhenian plate, and explains interaction between this wedge and the surrounding asthenosphere. The bottom of the wedge (the top of the Ionian slab) is constrained by the Benioff seismogenic zone¹². Its top (the base of the Tyrrhenian plate), as well as the base of the neighbouring Adriatic and African plates, is inferred from buoyancy considerations. Given a crustal structure, and considering crustal buoyancy, we can calculate the thickness of the mantle lithosphere that matches the total buoyancy of the lithosphere to the observed surface elevation^{13,14}. In practice, we remove the contribution of the crust to the observed surface elevation, and use the residual topography as a proxy for the thickness of the mantle lithosphere by assuming local isostasy¹⁰.

This calculation is demonstrated in the two cross-sections of Fig. 3: parallel and perpendicular to the direction of subduction. The black curve on the top of each cross-section describes the

contribution of the crust to the observed topography (H_c in ref. 10); this crustal contribution H_c is calculated from crustal thickness (compiled from the literature^{15–17}) and crustal mean density (estimated from seismic profiles). The red curve describes the residual topography (H_{ml} in ref. 10), from which we calculate the thickness of the mantle lithosphere, using $3,250 \text{ kg m}^{-3}$ and $3,200 \text{ kg m}^{-3}$ for the respective densities of the mantle lithosphere and the asthenosphere.

Because this procedure is based on the assumption that the lithosphere freely floats on the asthenosphere (that is, the system is in isostatic equilibrium), it does not apply to the region where the Ionian plate bends down towards the subduction zone. Here the slab-pull force is expressed by abnormally low residual topography (see red curve in section AA' southeast of Calabria) that cannot be interpreted in terms of mantle lithosphere thickness. In section AA' we schematically plotted the base of the Ionian plate with a dashed line, assuming that it is 125 km thick, but this is not constrained by data. More details about this method and its application to subduction zones, and to the central Mediterranean in particular are available elsewhere^{10,18}.

Section AA' shows that the Tyrrhenian plate is thin under the Vavilov basin and in the entire region between the Marsili basin and Calabria. The floors of the Vavilov and Marsili backarc basins are young oceanic crust¹⁹, and it is therefore not surprising that the lithosphere there is thin and the asthenosphere is high. However, upwelling of the asthenosphere closer to the plate boundary is unusual. In most cases, the upwelled asthenosphere reaches only the arc, and the crust of the forearc stays coupled to the subducting slab¹⁰. In this case, the asthenosphere has penetrated farther along the plate contact and reached the base of the crust under Calabria.

Section BB', taken perpendicular to section AA' at this point, reveals that the top of the Ionian slab under Calabria (inferred from the Benioff zone¹²) is deep enough to create a lateral 'window' of asthenosphere—a window that allows material to flow from under the African plate into the wedge between the base of the Tyrrhenian plate and the top of the Ionian slab. The direction of this flow, we suggest, is governed by the motion of the Ionian slab, which creates low pressure under Calabria. This flow is associated with significant upwelling from the base of the African lithosphere (~100 km deep) to the base of the crust under Calabria (~20 km deep). The upwelled, and hence decompressed, material partially melts, and the segregated magma finds its way to the surface through the faulted crust of Sicily, at the junction of the African plate with the

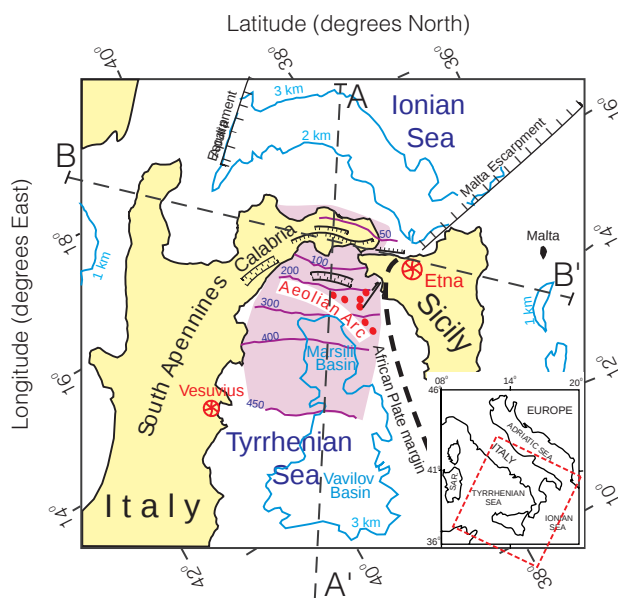


Figure 1 Map of the south Tyrrhenian subduction zone. The Ionian microplate, between the Apulia and Malta escarpments, subducts under the Tyrrhenian microplate, as indicated by the purple area (purple contours represent the depth to the top of the subducted slab, after Fregoli *et al.*¹²). The Aeolian islands (red dots) are a part of a volcanic arc that also includes several seamounts not shown here; the Vavilov and Marsili basins with their young oceanic crust (8–4.5 Myr ago and <2 Myr ago, respectively) are two backarc basins, and Calabria is the uplifted forearc region. No trench is observed east of Calabria due to a thick sedimentary fill. The normal faults and the newly developing rift zone between Calabria and the Aeolian islands are simplified from ref. 21. Lines marked as AA' and BB' show location of cross-sections in Figs 2 and 3. Inset: a regional location map of the central Mediterranean showing the studied area (red frame).

Ionian and Tyrrhenian microplates. It is at this triple junction that Mount Etna formed mostly during the last 200 kyr (ref. 11).

The existence of asthenospheric material under Calabria's crust explains not only the formation of Mount Etna, but also the recent uplift history of Calabria and the Aeolian arc¹⁰. Using the ages of uplifted marine terraces, Westaway²⁰ showed that Calabria began uplifting about 700 kyr ago. At about the same time, magmatic activity in the Aeolian islands greatly increased and also began to propagate eastward along a line connecting the Salina, Lipari and Volcano islands²¹ (the three islands on the southeast corner of the Aeolian arc—Fig. 1). Normal faulting and initiation of a new rift zone between Calabria and the Aeolian islands also took place at this time²¹ (Fig. 1).

We suggest that the increased magmatism at the arc, the uplifting of the forearc, and the initiation of a rift zone between these two regions, are all related to the rollback motion of the Ionian slab, which pulls the edge of the overriding plate and extends it. About 700–500 kyr go, the now-released top of the slab began to decouple from under Calabria and sink into the mantle. As a result, Calabria

began to rebound so that a gap filled by viscous material opened between the plates (section AA'). Initially, we suggest, the gap was filled by asthenospheric material that came from the west, passing under the arc region and enhancing the volcanic activity there. Later, as the slab sank deeper, a sideways flow from under the African plate was produced (section BB'), and the upwelling associated with this flow led to the formation of Mount Etna, as explained above.

Section AA' in Fig. 3 shows that in order to restore the slab to its position at about 500 kyr ago, and to close the gap between its top and the crust of Calabria, it is sufficient to raise it by ~70 km without significantly changing its shape (dotted line). The vertical sinking of the slab, as opposed to a change in its dip, is not only easier hydrodynamically, but would also explain why the top of the slab under the Aeolian islands is now deeper (200–250 km) than in most island arcs (~150 km, ref. 22). This, we believe, is a transient configuration following an episode of rapid slab detachment.

The sideways flow of asthenosphere associated with this episode is consistent with an earlier idea^{6,7} that the strongest suction

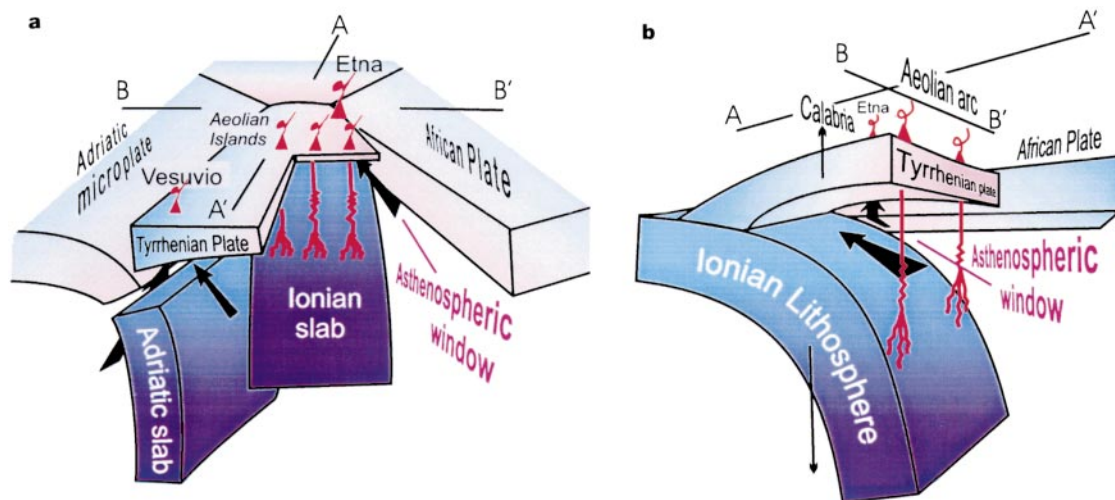


Figure 2 Three-dimensional sketch of the south Tyrrhenian subduction zone. Red lines represent magma rising from the top of the slab to the arc. Black arrows represent local patterns of mantle flow driven by slab motion. Notice that Etna is located outside the Aeolian arc and above a sideways mantle flow coming from under the African plate. **a**, A view toward the southeast from a point above the Tyrrhenian Sea. **b**, A view south and

upward from a point under the Tyrrhenian plate. This perspective illustrates the exceptional state of the south Tyrrhenian subduction zone, where asthenospheric material has penetrated high along the plate contact and reached the base of the crust in the forearc region. Only about 500 kyr ago this asthenospheric wedge did not exist, and the crust of Calabria was coupled to the top of the slab.

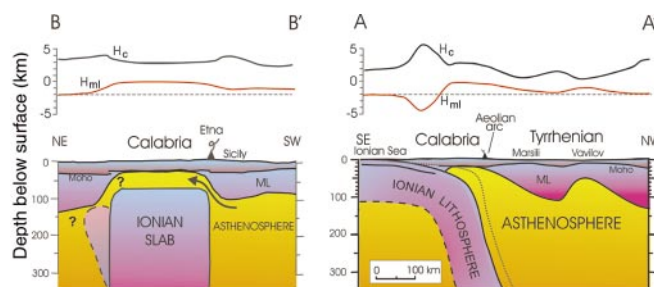


Figure 3 Cross-section parallel (AA') and perpendicular (BB') to the direction of subduction in the south Tyrrhenian subduction zone. Section AA' (right-hand panels) shows the thinning of Tyrrhenian lithosphere, and the penetration of asthenosphere between the subducting and overriding plates. The top of the Ionian slab is inferred from the Benioff zone¹², and the bottom is plotted schematically (dashed line) assuming a constant thickness of 125 km. The dotted line represents the location of the top of the slab before decoupling ~500 kyr ago. Section BB' (left-hand panels) shows the asthenospheric window that allows flow from under the African plate to the base of the crust of Calabria. We suggest that decompression associated with this flow is the cause of the voluminous melting under Mount Etna. The pattern of mantle flow in the northern end of the subduction zone is unclear (question marks), owing to the unknown geometry of the Adriatic slab (dashed line, refs 25–28). These cross-sections are based on the separation of topography into its crustal (H_c , in black) and mantle lithosphere (H_{ml} , in red) contributions. The thickness of the crust is known, and the thickness of the mantle lithosphere is calculated from H_{ml} . The dashed line at $H_{ml} = -2$ km represents the normal contribution to the Earth's topography of a mantle lithosphere 90–110 km thick. Here, significant variations from this reference value are observed. Details are available elsewhere^{10,18}.

associated with slab rollback is produced at the corner between the plates. It also matches the observation that narrow slabs roll back relatively quickly, because the flow around their edges significantly reduces the corner suction^{6,7}. The Ionian slab, which is narrower than it is long (Fig. 1), thus appears to be an excellent example of this process.

The truncated shape of this slab at its southern edge is clearly indicated by the abrupt cessation of seismic activity approximately at the north coast of Sicily. About 15 Myr ago this edge was torn away from another slab fragment, now detected by seismic tomography under the coasts of Algeria²³. The Ionian fragment rolled back to its present position, opening the Tyrrhenian backarc basin, whereas its other fragment sank in the mantle under the coasts of Algeria²⁴.

The shape of the northern edge of the Ionian slab is less clear. Earthquake distribution and tomographic studies seem to indicate that the Ionian slab is disconnected from the nearby Adriatic slab^{25,26} and that the Adriatic slab is decoupled from its mother plate, Adria^{26–28}. In Fig. 2a we suggest that here, the sinking Adriatic plate creates a local flow in the asthenosphere that produced the Quaternary magmatism at Mount Vesuvius and nearby. However, the exact subsurface geometry at this triple junction, between the Ionian, Tyrrhenian and Adriatic microplates, is yet to be explored. □

Received 15 April; accepted 16 August 1999.

- Barberi, F. *et al.* Evolution of a section of the African-Europe plate boundary: paleomagnetic and volcanological evidence from Sicily. *Earth Planet. Sci. Lett.* **21**, 269–276 (1974).
- Condomines, M., Tanguy, J. C., Kieffer, G. & Allegre, C. J. Magmatic evolution of a volcano studied by ²³⁸Th, ²³⁵U disequilibrium and trace elements systematics: The Etna case. *Geochim. Cosmochim. Acta* **46**, 1397–1416 (1982).
- Tanguy, J. C. Tholeiitic basalt magmatism of Mount Etna and its relations with alkaline series. *Contrib. Mineral. Petrol.* **66**, 51–67 (1978).
- Allard, P. *et al.* Mantle-derived helium and carbon in groundwater and gases of Mount Etna, Italy. *Earth Planet. Sci. Lett.* **148**, 501–516 (1997).
- Allard, P. *et al.* Eruptive and diffuse emissions of CO₂ from Mount Etna. *Nature* **351**, 387–391 (1991).
- Nur, A., Dvorkin, J., Mavko, G. & Ben Avraham, Z. Speculations on the origin and fate of backarc basins. *Ann. Geofis.* **36**, 155–163 (1991).
- Dvorkin, J., Nur, A., Mavko, G. & Ben Avraham, Z. Narrow subducting slabs and the origin of backarc basins. *Tectonophysics* **227**, 63–79 (1993).
- Wendt, J. I., Regelous, M., Collerson, K. D. & Ewart, A. Evidence for a contribution from two mantle plumes to the island arc lavas from northern Tonga. *Geology* **25**, 611–614 (1997).
- Turner, S. & Howkesworth, C. Using geochemistry to map mantle flow beneath the Lau basin. *Geology* **26**, 1019–1022 (1998).
- Gvirtzman, Z. & Nur, A. Plate detachment, asthenosphere upwelling, and topography across subduction zones. *Geology* **27**, 563–566.
- Monaco, C., Tappinier, P., Tortorici, L. & Gillot, P. Y. Late quaternary slip rates on the Acireale-Piedimonte normal fault and tectonic origin of Mt. Etna (Sicily). *Earth Planet. Sci. Lett.* **147**, 125–139 (1997).
- Frepoli, A., Selvaggi, G., Chiarabba, C. & Amato, A. State of stress in southern Tyrrhenian subduction zone from fault-plane solutions. *Geophys. J. Int.* **125**, 879–891 (1996).
- Lachenbruch, A. H. & Morgan, P. Continental extension, magmatism and elevation: Formal relations and rules of thumb. *Tectonophysics* **174**, 39–62 (1990).
- Jones, H. J., Unruh, J. R. & Sonder, L. J. The role of gravitational potential energy in active deformation in the southwestern United States. *Nature* **381**, 37–41 (1996).
- Steinmetz, L., Ferrucci, F., Hirn, C. A., Morelli, C. & Nicolich, R. A 550 km long Moho traverse in the Tyrrhenian Sea from O.B.S. recorded P_n waves. *Geophys. Res. Lett.* **10**, 428–431 (1983).
- Scarcia, S., Lozej, A. & Cassinis, R. Crustal structures of the Ligurian, Tyrrhenian, and Ionian seas and adjacent onshore areas interpreted from wide-angle seismic profiles. *Boll. Geofis. Teor. Applic.* **36**, 5–19 (1994).
- Morelli, C. Current knowledge on the crustal properties of Italy. *Ann. Geofis.* **37**, 1113–1130 (1997).
- Gvirtzman, Z. & Nur, A. Residual topography, lithospheric structure and sunken slabs in the central Mediterranean. *Earth Planet. Sci. Lett.* (submitted).
- Kastens, K. *et al.* ODP Leg 107 in the Tyrrhenian Sea: Insight into passive and backarc basin evolution. *Geol. Soc. Am. Bull.* **100**, 1140–1156 (1988).
- Westaway, R. Quaternary uplift of south Italy. *J. Geophys. Res.* **98**, 21741–21772 (1993).
- Mazzuoli, R., Tortorici, L. & Ventura, G. Oblique rifting in Salina, Lipari, and Vulcano islands (Aeolian islands, southern Italy). *Terra Nova* **7**, 444–452 (1995).
- Isacks, B. L. & Barazangi, M. In *Island Arcs, Deep Sea Trenches, and Backarc Basins* (eds Talawani, M. and Pitman, W. C.) 99–114 (Maurice Ewing Ser. 1, American Geophysical Union, Washington DC, 1977).
- Bijwaard, H., Spakman, W. & Engdahl, E. R. Closing the gap between regional and global travel time tomography. *J. Geophys. Res.* **103**, 30053–30078 (1998).
- Carminati, E., Wortel, M. J. R., Spakman, W. & Sabadina, R. The role of slab detachment processes in the opening of the western-central Mediterranean basin: some geological and geophysical evidence. *Earth Planet. Sci. Lett.* **160**, 651–665 (1998).
- Mele, G. High frequency wave propagation from mantle earthquakes in the Tyrrhenian Sea: New

constraints for the geometry of the south Tyrrhenian subduction zone. *Geophys. Res. Lett.* **25**, 2877–2880 (1998).

- Amato, A., Alessandrini, B., Cimini, G. B., Frepoli, A. & Selvaggi, G. Active and remnant subducted slabs beneath Italy: evidence from seismic tomography and seismicity. *Ann. Geofis.* **36**, 201–214 (1993).
- Wortel, M. J. R. & Spakman, W. Structure and dynamics of subducted lithosphere in the Mediterranean region. *Proc. Koninklijke Ned. Acad. Wetensch.* **95**, 325–347 (1992).
- Spakman, W., van de Lee, S. & van der Hilst, R. Travel time tomography of the European-Mediterranean mantle down to 1400 km. *Phys. Earth Planet. Inter.* **79**, 3–74 (1993).

Correspondence and requests for materials should be addressed to Z.G. (e-mail: zohar@vms.huji.ac.il).

Decline in Mesozoic reef-building sponges explained by silicon limitation

Manuel Maldonado, M. Carmen Carmona, María J. Uriz & Antonio Cruzado

Centro de Estudios Avanzados de Blanes (CSIC), Camino de Santa Barbara s/n, Blanes 17300, Girona, Spain

Several unrelated clades of siliceous sponges proliferated on the shelves of the Jurassic Tethys Sea, becoming prominent builders in reefs and near-shore mounds^{1–4}. Many of these builders are characterized by massive, rock-like skeletons made of spicules with a characteristic terminal hypersilicification^{4,5}. Such hypertrophied spicules are generically known as desmas, irrespective of their phylogenetic origin⁵. Desma-bearing sponges virtually disappeared from reefs and other neritic environments during the Cretaceous and the Early Tertiary^{1,2,4,6}, but have subsisted in relict populations in deeper, bathyal waters^{5,7,8}. The causes of the decline and bathymetric shift of these sponges remain obscure. Here we show experimentally that the concentration of silicic acid in seawater modulates the phenotypic expression of the various spicule types genetically available in a sponge species. We also show that the concentration of this nutrient in Recent surface waters is insufficient for this species to secrete its desmas. These findings indicate that silicon limitation, probably aggravated in shallow waters by the diatom burst around the Cretaceous–Tertiary boundary^{9,10}, may have forced neritic sponges with desmas to either lighten their skeletons or move to deeper, silicon-rich environments.

Silicon (Si) is one of the most abundant elements on earth, but only its soluble form, as monomeric, undissociated silicic acid, Si(OH)₄, is biologically assimilable. Several groups of marine organisms, such as diatoms, radiolarians, choanoflagellates and sponges, take up Si(OH)₄ from seawater to build their opal (amorphous hydrated silica) skeletons. Biological consumption of Si(OH)₄ renders present-day oceans undersaturated with this nutrient. Yet deep waters are relatively rich in Si(OH)₄, with concentrations varying from 10 to 180 μM, whereas surface waters are notably poorer, with a mean concentration of less than 3 μM^{11,12}.

Diatoms are responsible for the depletion of Si(OH)₄ in the photic zone and largely control the cycling of Si in Recent oceans¹³. Radiolarians are quantitatively important only in the Pacific equatorial belt¹⁴, whereas siliceous sponges and silicoflagellates are considered unimportant on a global scale¹⁰. However, there is sedimentological evidence that sponges once controlled the Si cycling in neritic environments¹⁰. Indeed, desma-bearing sponges and hexactinellids, two unrelated groups with substantial siliceous

skeletons, proliferated in neritic assemblages and became prominent builders in reefs and near-shore mounds during the Upper Jurassic^{1–4}. We show that the encrusting demosponge *Crambe crambe* (Schmidt), one of the most common macroinvertebrates in the western Mediterranean sublittoral¹⁵, provides some clues to understand why sponges with massive silica skeletons declined in neritic environments.

In most natural populations, skeletons of *C. crambe* contain only a single spicule type, small needles called styles¹⁶; however, specimens are occasionally found in which the skeleton is supplemented by large styles, scarce C-shaped spicules (called isochelae) or scarce aster-like desmas¹⁶. This variability has puzzled sponge biologists, as the number of spicule types secreted by a species is a cornerstone character in sponge taxonomy; it is assumed that the spicule set does

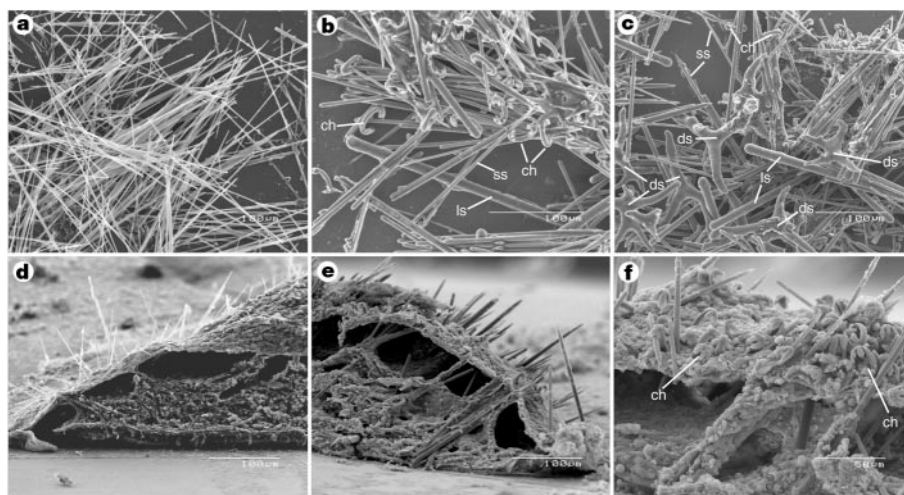


Figure 1 Skeletons produced under different concentrations of $\text{Si}(\text{OH})_4$. **a**, Skeleton of a control sponge built up by a single, needle-like spicule type called style. **b**, Skeleton of a sponge reared in $30 \mu\text{M}$ $\text{Si}(\text{OH})_4$ showing three spicule types: small styles (ss); large styles (ls); and small C-shaped spicules called isochelae (ch). Note differences in size and robustness between styles of control sponges and styles of sponges reared in Si-enriched

seawater. **c**, Skeleton of a sponge reared in $100 \mu\text{M}$ $\text{Si}(\text{OH})_4$ showing four spicule types: styles in two size categories (ss, ls); isochelae (ch); and aster-like desmas (ds). **d, e**, Skeletal arrangements seen by fracturing sponges reared in controls and in $30 \mu\text{M}$ $\text{Si}(\text{OH})_4$, respectively. **f**, Ectosome of a sponge reared in $30 \mu\text{M}$ $\text{Si}(\text{OH})_4$ showing great abundance of protruding isochelae (ch).

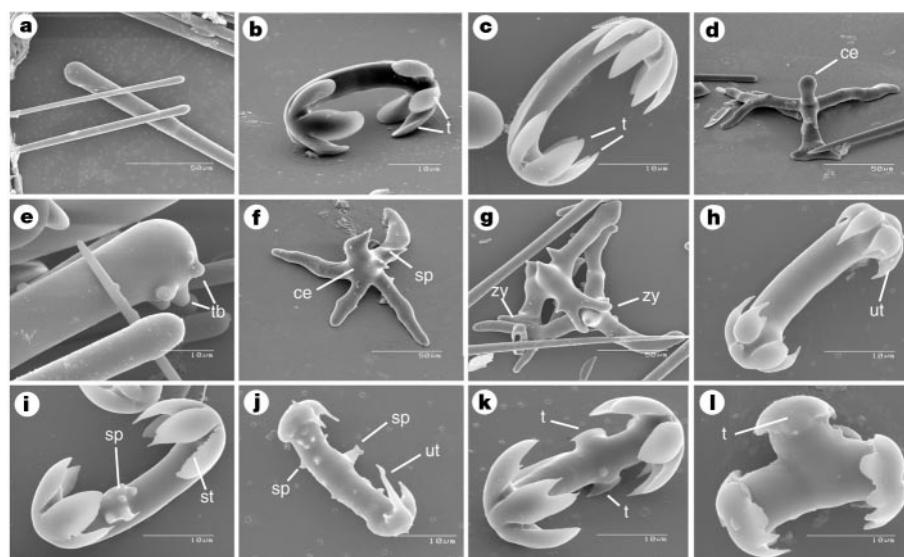


Figure 2 Spicules produced under two experimental levels of Si supply. **a–d**, Spicules produced in $30 \mu\text{M}$ $\text{Si}(\text{OH})_4$. **a**, Styles in two size categories. **b, c**, Isochelae showing intraspecific variability in the shape and number of teeth (t), traits assumed to have taxonomic value in some sponge groups. **d**, An aster-like desma with an undeveloped centrum (ce). **e–l**, Spicules produced in $100 \mu\text{M}$ $\text{Si}(\text{OH})_4$. **e**, Small and large styles, the latter with a tuberosity (tb), a trait of presumed taxonomic value. **f**, Aster-like

desma showing a well-developed centrum (ce) with three large spines (sp). **g**, Desmas still in contact after treatment with nitric acid, indicating some degree of fusion at the contact points; zy, zygos. **h–l**, Common, malformed isochelae showing teeth reduced to unguis (ut) (**h**), serrated teeth (st) and spines (sp) on the shaft (**i**), reduced teeth (ut) and spines (sp) on the shaft (**j**), and teeth (t) on the shaft (**k–l**).

not vary within species.

We examined the production of spicules by newly settled sponges maintained at 3 concentrations of $\text{Si}(\text{OH})_4$ for 14 weeks (August to November 1998). One set of 36 sponges was cultured in seawater from their natural habitat. Concentration of $\text{Si}(\text{OH})_4$ in these controls ranged from 0.03 to 4.5 μM (mean $\pm$ s.e. = $0.741 \pm 0.133 \mu\text{M}$). Two other sets were reared at elevated concentrations of $\text{Si}(\text{OH})_4$, referred to hereafter as Si-medium ($30.235 \pm 0.287 \mu\text{M}$) and Si-high ($100.041 \pm 0.760 \mu\text{M}$) treatments.

All sponges reared in control cultures formed only a single spicule type: small styles (Fig. 1a, d). In contrast, all sponges reared in Si-medium conditions secreted at least three spicule types in great abundance: small styles, large styles and 28–33- μm long isochelae (Figs 1b, e–f, 2a–c). Fifty per cent of these sponges also formed a small number (1–5) of a fourth spicule type: aster-like desmas (Fig. 2d). In Si-high conditions, all sponges produced all four spicule types in great abundance (Figs 1c, 2e–l). Two of these sponges also secreted a fifth category: 8–10- μm long isochelae (spicule not shown). Notably, all sponges used in the experiments were from a large population of *C. crambe* studied for more than 15 years, in which we never found large styles, isochelae or desmas.

The concentration of $\text{Si}(\text{OH})_4$ had a striking effect on both the size and shape of spicules (Figs 1a–c, 2). In Si-high conditions, a few isochelae formed spines and tooth-like structures at the middle of the shaft (Fig. 2k, l), which resembled spicules from the fossil record^{4,17}. Differences in $\text{Si}(\text{OH})_4$ concentration also affected its uptake, as determined by analyses during the first five weeks of the experiment. Consumption curves of $\text{Si}(\text{OH})_4$ were parallel in all three treatments, with a peak immediately after metamorphosis (Fig. 3). Sponges took up $\text{Si}(\text{OH})_4$ more efficiently in Si-medium than in Si-high conditions, indicating that silica transport into skeletal secretory cells (sclerocytes) may be inhibited at unusually high concentrations. This may explain the production of numerous malformed isochelae in Si-high conditions (Fig. 2h–l). More significantly, $\text{Si}(\text{OH})_4$ uptake was higher in the two Si-enriched treatments than in controls (Fig. 3), which indicates that *C. crambe* is limited by Si availability in coastal waters.

Our results show that *C. crambe* is genetically capable of producing spicule types that are not normally found in natural populations. Limited silicic acid in coastal waters appears to be the reason for this phenotypic skeletal inhibition, which affects entire populations of thousands of individuals. Our results also indicate that, in contrast to what is believed, each spicule type may be secreted by a specific sclerocyte type, which is activated by a particular $\text{Si}(\text{OH})_4$ concentration threshold. Significantly, aster-like desmas are consistently formed only at concentrations of $\text{Si}(\text{OH})_4$ (100 μM) that are not found in most surface waters today. Sponges with aster-like desmas were fairly common in Jurassic–Cretaceous assemblages,

but all these are now extinct, except the sublittoral genus *Crambe* and, perhaps, the monotypic, deep-water genus *Vetulina*^{4,16,18}.

The possibility that $\text{Si}(\text{OH})_4$ concentration influenced the skeletal evolution of sponges is discussed elsewhere^{4,19}. Two peaks of volcanic activity in the Triassic (Fassanian and Sevatian), which probably increased the amount of dissolved Si in seawater, may have favoured the radiation and proliferation of sponges with massive siliceous skeletons during the Jurassic^{4,19}. The hypothesis that seawater was saturated with dissolved Si at that time is also supported by the robust skeletons of fused spicules and state of preservation of fossils^{9,20}. The study of chert beds indicates that domination of the Si cycle by diatoms was completed only in the Eocene, after a long period of expansion and radiation beginning during the Cretaceous²¹. The contribution of sponge spicules to neritic cherts began to decline concurrently with the radiation and proliferation of diatoms^{9,10}. The palaeontological record indicates that both desma-bearing and hexactinellid sponges had disappeared from neritic environments and penetrated the bathyal and abyssal zones by the Lower Tertiary^{1,4,6}. During the Cenozoic, siliceous spicules of neritic sponges became progressively less robust and fused spicules less frequent^{6,9}. The few Recent hexactinellids still found at shallow depths occur either in microhabitats with enhanced Si supply²² or at high latitudes^{23,24}, where the concentration of $\text{Si}(\text{OH})_4$ in surface water is higher. There is also correlative evidence that the expansion of diatoms reduced both diversity and abundance of other Si-dependent organisms, such as radiolarians²⁵.

The hypothesis that low $\text{Si}(\text{OH})_4$ concentration in shallow waters might limit sponges has been disregarded^{7,26,27}, as the Si-uptake system of sponges is expected to have adapted to the decline in the availability of this nutrient. In contrast, our results indicate that Si concentration is a limiting factor on shallow-water sponges and that this pressure may be strong enough to shape their skeletons and restrict their bathymetric distributions. Chronic Si limitation is also supported by the only study on Si uptake in a sublittoral sponge, *Halichondria panicea*. This showed Michaelis–Menten-uptake kinetics with saturation at $\text{Si}(\text{OH})_4$ concentrations far higher than the maximum available in the natural habitat of the species²⁸. We propose that chronic limitation may be responsible not only for much of the unexplained skeletal variability that puzzles sponge taxonomists, but also for much of the skeletal change observed in shallow-water sponges during the Cenozoic. In which sponge groups the 'skeletal change' represents evolution, and in which it merely reflects phenotypic inhibition of certain spicule types remain unclear. Therefore, we cannot rule out the possibility that by rearing more sponges in seawater saturated with silicic acid some apparently extinct species or spicules may 'come back to life' for taxonomists. □

Methods

The sponges used in the experiment were grown in the laboratory from free-swimming

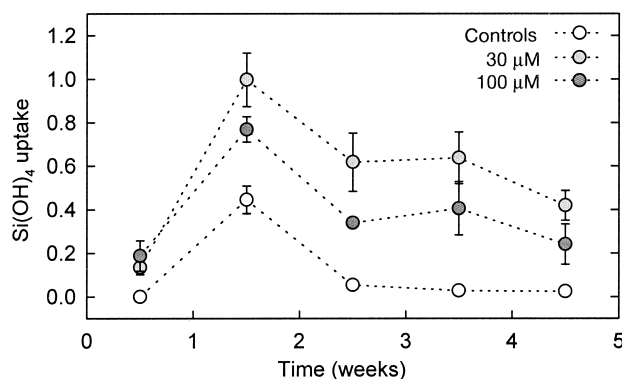


Figure 3 Consumption (mean $\pm$ s.e.) of $\text{Si}(\text{OH})_4$ by sponges (μmol per individual per week) as a function of substrate concentration and time during the first five weeks of the experiment.

larvae collected randomly in the field (2° 48.12' N, 41° 40.33' E) by SCUBA. Between 5 and 10 juveniles were recruited successfully in each of 15, 1 l polystyrene containers ($n = 15$), the bottom of which was covered with an acetate sheet that served as substratum for sponge attachment. Containers were then randomly distributed in 3 groups, and sponges in each group were reared for 14 weeks in 3 different concentrations of $\text{Si}(\text{OH})_4$: 0.741 ± 0.133 , 30.235 ± 0.287 and $100.041 \pm 0.760 \mu\text{M}$ (mean $\pm$ s.e.). All cultures were prepared using $0.22 \mu\text{M}$ polycarbonate-filtered seawater, which was collected from the sponge habitat, handled according to standard methods to prevent Si contamination²⁹ and enriched in dissolved silica, when treatments required, by using Na_2SiF_6 . During the experiment, all sponges were fed by weekly addition of 2 ml of a bacterial culture ($40\text{--}60 \times 10^6$ bacteria ml^{-1}) to each container³⁰. The seawater was replaced weekly, with regeneration of initial food and $\text{Si}(\text{OH})_4$ levels. The concentration of $\text{Si}(\text{OH})_4$ in cultures was determined on 3 replicates of 1 ml seawater samples per container by using a Bran-Luebbe TRAACS 2000 nutrient autoanalyser. After week 5, the accidental contamination of some culture containers by diatoms rendered subsequent estimates of Si uptake by sponges unreliable, so we discarded them for the study.

For the study of the skeleton, sponges were treated according to standard methods³⁰ and examined in a Hitachi S-2300 scanning electron microscope (SEM).

Received 21 April; accepted 16 August 1999.

- Hartman, W. D., Wendt, J. W. & Wiedenmayer, F. Living and fossil sponges. Notes for a short course. *Sedimentaria* **8**, 1–274 (1980).
- Ghiold, J. The sponges that spanned Europe. *New Scient.* **129**, 58–62 (1991).
- Leinfelder, R. R. Upper Jurassic reef types and controlling factors. *Profil* **5**, 1–45 (1993).
- Wiedenmayer, F. Contributions to the knowledge of post-Paleozoic neritic and archibenthal sponges (Porifera). *Schweiz. Paläont. Abh.* **116**, 1–147 (1994).
- Lévi, C. in *Fossil and Recent Sponges* (eds Reiter, J. & Keupp, H.) 72–82 (Springer, New York, 1991).
- Moret, L. Contribution à l'étude des spongiaires siliceux du Miocene de l'Algérie. *Mém. Soc. Géol. Fr.* **1**, 1–27 (1924).
- Vaculet, J. Indications de profondeur données par les Spongiaires dans les milieux benthiques actuels. *Géol. Méditerranée* **15**, 13–26 (1988).
- Maldonado, M. & Young, C. M. Bathymetric patterns of sponge distribution on the Bahamian slope. *Deep-Sea Res.* **143**, 897–915 (1996).
- Lowenstam, H. A. & Weiner, S. *On Biomineralization* (Oxford Univ., Oxford, 1989).
- Maliva, R. G., Knoll, A. H. & Siever, R. Secular change in chert distribution: a reflection of evolving biological participation in the silica cycle. *Palaios* **4**, 519–532 (1989).
- Nelson, D. M., Tréguer, P., Brzezinski, M. A., Leynaert, A. & Quéguiner, B. Production and dissolution of biogenic silica in the ocean: revised global estimates, comparison with regional data and relationship to biogenic sedimentation. *Glob. Biochem. Cycles* **9**, 359–372 (1995).
- Tréguer, P. et al. The silica balance in the world ocean: a reestimate. *Science* **268**, 375–379 (1995).
- Calvert, S. E. in *Silicon Geochemistry and Biogeochemistry* (ed. Aston, S. R.) 143–186 (Academic, London, 1983).
- Lisitzyn, A. P. Sedimentation in the world ocean. *Soc. Econ. Palaeon. Mineral. Spec. Pub.* **17**, 1–218 (1972).
- Weinberg, S. *Découvrir la Méditerranée* (Éditions Nathan, Paris, 1993).
- Maldonado, M. & Uriz, M. J. Skeletal morphology of two controversial poecilosclerid genera (Porifera, Demospongiae): *Discorhabdella* and *Crambe*. *Helgoländer Meeresunters.* **50**, 369–390 (1996).
- Hinde, G. J. & Holmes, W. M. On the sponge remains in the Lower Tertiary strata near Oamaru, New Zealand. *J. Linn. Soc. Zool.* **24**, 177–262 (1892).
- Kelly-Borges, M. & Pomponi, S. A. Phylogeny and classification of lithistid sponges (Porifera: Demospongiae): a preliminary assessment using ribosomal DNA sequence comparisons. *Mol. Mar. Biol. Biotech.* **3**, 87–103 (1994).
- Mostler, H. Poriferenspiculae der alpinen Trias. *Geol. Paläont. Mitt. Innsbruck* **6**, 1–42 (1976).
- Palmer, T. J. & Fürsich, F. T. Ecology of sponge reefs from the Upper Bathonian of Normandy. *Palaeontology* **24**, 1–23 (1981).
- Burckle, L. H. in *Introduction to Marine Micropaleontology* (eds Haq, B. U. & Boersma, A.) 245–266 (Elsevier, Amsterdam, 1978).
- Austin, B. Underwater birdwatching. *Canadian Tech. Rep. Hydro. Ocean. Sci.* **38**, 83–90 (1984).
- Koltun, V. M. in *The Biology of the Porifera* (ed. Fry, W. G.) 285–297 (Academic, London, 1970).
- Tabachnick, K. R. in *Sponges in Time and Space* (eds van Soest, R. M. W., van Kempen, T. M. G. & Brackman, J. C.) 225–232 (A. A. Balkema, Rotterdam, 1994).
- Harper, H. E. & Knoll, A. H. Silica, diatoms and Cenozoic radiolarian evolution. *Geology* **3**, 175–177 (1975).
- Hartman, W. D. in *Silicon and Siliceous Structures in Biological Systems* (eds Simpson, T. L. & Volcani, B. E.) 453–493 (Springer Verlag, New York, 1981).
- Pisera, A. Upper Jurassic siliceous sponges from Swabian Alb: taxonomy and paleoecology. *Palaeont. Pol.* **57**, 1–216 (1997).
- Reincke, T. & Barthel, D. Silica uptake kinetics of *Halichondria panicea* in Kiel Bight. *Mar. Biol.* **129**, 591–593 (1997).
- Grasshoff, K., Ehrhardt, M. & Kremling, K. *Methods of Seawater Analysis* (Verlag Chemie, Weinheim, 1983).
- Maldonado, M. & Uriz, M. J. An experimental approach to the ecological significance of microhabitat-scale movement in an encrusting sponge. *Mar. Ecol. Prog. Ser.* **185**, 239–255 (1999).

Acknowledgements

We thank S. Pla for help with nutrient analyses, technicians of the Servicio de Microscopia for help with SEM, and E. Ballesteros, C. M. Young, A. Pisera and R. Rycroft for comments on the manuscript.

Correspondence and requests for materials should be addressed to M.M. (e-mail: maldonado@ceab.csic.es).

Learning the parts of objects by non-negative matrix factorization

Daniel D. Lee* & H. Sebastian Seung*†

* Bell Laboratories, Lucent Technologies, Murray Hill, New Jersey 07974, USA

† Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

Is perception of the whole based on perception of its parts? There is psychological¹ and physiological^{2,3} evidence for parts-based representations in the brain, and certain computational theories of object recognition rely on such representations^{4,5}. But little is known about how brains or computers might learn the parts of objects. Here we demonstrate an algorithm for non-negative matrix factorization that is able to learn parts of faces and semantic features of text. This is in contrast to other methods, such as principal components analysis and vector quantization, that learn holistic, not parts-based, representations. Non-negative matrix factorization is distinguished from the other methods by its use of non-negativity constraints. These constraints lead to a parts-based representation because they allow only additive, not subtractive, combinations. When non-negative matrix factorization is implemented as a neural network, parts-based representations emerge by virtue of two properties: the firing rates of neurons are never negative and synaptic strengths do not change sign.

We have applied non-negative matrix factorization (NMF), together with principal components analysis (PCA) and vector quantization (VQ), to a database of facial images. As shown in Fig. 1, all three methods learn to represent a face as a linear combination of basis images, but with qualitatively different results. VQ discovers a basis consisting of prototypes, each of which is a whole face. The basis images for PCA are 'eigenfaces', some of which resemble distorted versions of whole faces⁶. The NMF basis is radically different: its images are localized features that correspond better with intuitive notions of the parts of faces.

How does NMF learn such a representation, so different from the holistic representations of PCA and VQ? To answer this question, it is helpful to describe the three methods in a matrix factorization framework. The image database is regarded as an $n \times m$ matrix V , each column of which contains n non-negative pixel values of one of the m facial images. Then all three methods construct approximate factorizations of the form $V \approx WH$, or

$$V_{i\mu} \approx (WH)_{i\mu} = \sum_{a=1}^r W_{ia} H_{a\mu} \quad (1)$$

The r columns of W are called basis images. Each column of H is called an encoding and is in one-to-one correspondence with a face in V . An encoding consists of the coefficients by which a face is represented with a linear combination of basis images. The dimensions of the matrix factors W and H are $n \times r$ and $r \times m$, respectively. The rank r of the factorization is generally chosen so that $(n + m)r < nm$, and the product WH can be regarded as a compressed form of the data in V .

The differences between PCA, VQ and NMF arise from different constraints imposed on the matrix factors W and H . In VQ, each column of H is constrained to be a unary vector, with one element equal to unity and the other elements equal to zero. In other words, every face (column of V) is approximated by a single basis image (column of W) in the factorization $V \approx WH$. Such a unary encoding for a particular face is shown next to the VQ basis in Fig. 1. This unary representation forces VQ to learn basis images that are prototypical faces.

PCA constrains the columns of W to be orthonormal and the rows of H to be orthogonal to each other. This relaxes the unary constraint of VQ, allowing a distributed representation in which each face is approximated by a linear combination of all the basis images, or eigenfaces⁶. A distributed encoding of a particular face is shown next to the eigenfaces in Fig. 1. Although eigenfaces have a statistical interpretation as the directions of largest variance, many of them do not have an obvious visual interpretation. This is because PCA allows the entries of W and H to be of arbitrary sign. As the eigenfaces are used in linear combinations that generally involve complex cancellations between positive and negative numbers, many individual eigenfaces lack intuitive meaning.

NMF does not allow negative entries in the matrix factors W and H . Unlike the unary constraint of VQ, these non-negativity constraints permit the combination of multiple basis images to represent a face. But only additive combinations are allowed, because the non-zero elements of W and H are all positive. In contrast to PCA, no subtractions can occur. For these reasons, the non-negativity constraints are compatible with the intuitive notion of combining parts to form a whole, which is how NMF learns a parts-based representation.

As can be seen from Fig. 1, the NMF basis and encodings contain a large fraction of vanishing coefficients, so both the basis images and image encodings are sparse. The basis images are sparse because they are non-global and contain several versions of mouths, noses and other facial parts, where the various versions are in different locations or forms. The variability of a whole face is generated by combining these different parts. Although all parts are used by at

least one face, any given face does not use all the available parts. This results in a sparsely distributed image encoding, in contrast to the unary encoding of VQ and the fully distributed PCA encoding⁷⁻⁹.

We implemented NMF with the update rules for W and H given in Fig. 2. Iteration of these update rules converges to a local maximum of the objective function

$$F = \sum_{i=1}^n \sum_{\mu=1}^m [V_{i\mu} \log(WH)_{i\mu} - (WH)_{i\mu}] \quad (2)$$

subject to the non-negativity constraints described above. This objective function can be derived by interpreting NMF as a method for constructing a probabilistic model of image generation. In this model, an image pixel $V_{i\mu}$ is generated by adding Poisson noise to the product $(WH)_{i\mu}$. The objective function in equation (2) is then related to the likelihood of generating the images in V from the basis W and encodings H .

The exact form of the objective function is not as crucial as the non-negativity constraints for the success of NMF in learning parts. A squared error objective function can be optimized with update rules for W and H different from those in Fig. 2 (refs 10, 11). These update rules yield results similar to those shown in Fig. 1, but have the technical disadvantage of requiring the adjustment of a parameter controlling the learning rate. This parameter is generally adjusted through trial and error, which can be a time-consuming process if the matrix V is very large. Therefore, the update rules described in Fig. 2 may be advantageous for applications involving large data-bases.

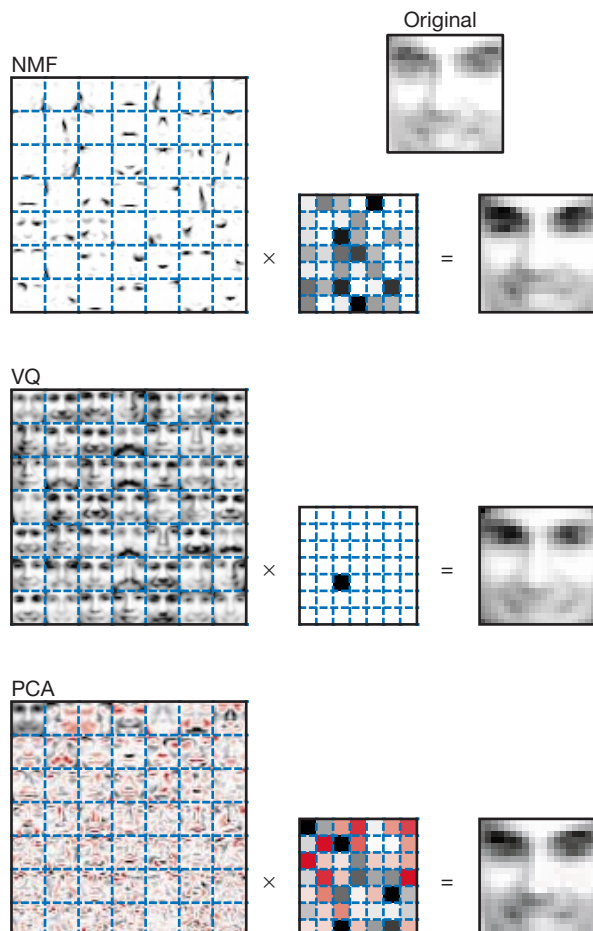


Figure 1 Non-negative matrix factorization (NMF) learns a parts-based representation of faces, whereas vector quantization (VQ) and principal components analysis (PCA) learn holistic representations. The three learning methods were applied to a database of $m = 2,429$ facial images, each consisting of $n = 19 \times 19$ pixels, and constituting an $n \times m$ matrix V . All three find approximate factorizations of the form $V \approx WH$, but with three different types of constraints on W and H , as described more fully in the main text and methods. As shown in the 7×7 montages, each method has learned a set of $r = 49$ basis images. Positive values are illustrated with black pixels and negative values with red pixels. A particular instance of a face, shown at top right, is approximately represented by a linear superposition of basis images. The coefficients of the linear superposition are shown next to each montage, in a 7×7 grid, and the resulting superpositions are shown on the other side of the equality sign. Unlike VQ and PCA, NMF learns to represent faces with a set of basis images resembling parts of faces.

It is helpful to visualize the dependencies between image pixels and encoding variables in the form of the network shown in Fig. 3. The top layer of nodes represents an encoding $h_1, \dots, h_r$ (column of H), and the bottom layer an image $v_1, \dots, v_n$ (column of V). The matrix element W_{ia} quantifies the amount of influence that the a th encoding variable h_a has on the i th image pixel v_i . A single encoding variable influences multiple image pixels, owing to the fan-out of connections from the encoding variable. Because of the non-negativity of W_{ia} , this influence is restricted to coactivation of image pixels. Intuitively, a parts-based representation should be learnable from observations of coactivation in V , as the image pixels belonging to the same part of the face are coactivated when that part is present. NMF learns by adapting W_{ia} to generate the appropriate coactivations.

The preceding description of NMF has been specialized to images, but the algorithm is actually applicable to a wide variety of problem domains. More generally, NMF is a method for modelling the generation of directly observable visible variables V from hidden variables H (refs 12, 13). Each hidden variable coactivates a subset of visible variables, or 'part'. Activation of a constellation of hidden variables combines these parts additively to generate a whole. Seen in this light, NMF has a very broad range of potential applications. We illustrate this versatility by applying NMF to a completely different problem, the semantic analysis of text documents.

For this application, a corpus of documents is summarized by a matrix V , where $V_{i\mu}$ is the number of times the i th word in the vocabulary appears in the μ th document¹⁴. These word counts can be regarded as a set of visible variables and modelled as being generated from an underlying set of hidden variables. Application of VQ, PCA or NMF involves finding the approximate factorization of this matrix $V \approx WH$ into a feature set W and hidden variables H , in the same way as was done for faces.

In the VQ factorization, a single hidden variable is active for each document. If the same hidden variable is active for a group of documents, they are semantically related, because they have similar frequencies of word occurrence. Consequently, the hidden variables are called semantic variables, and VQ is accordingly used for automatic semantic indexing of documents by topic¹⁴. Each column of W , or semantic feature, consists of the word frequencies for the corresponding semantic variable.

VQ allows only one semantic variable to be active, which prevents more than one topic from being attributed to a document. PCA would seem to be a solution to this problem, as it allows activation of multiple semantic variables. Although PCA has been successful in

certain linguistic tasks¹⁵, it generally results in semantic variables that are difficult to interpret, for much the same reason that the PCA representation of faces has no obvious visual interpretation. This is the result of two unrealistic aspects of the model: all semantic variables are used to represent each document; and negative values for semantic variables are allowed. Intuitively, it makes more sense for each document to be associated with some small subset of a large array of topics, rather than just one topic or all the topics. Because the sparsely distributed representation of NMF appears ideally suited for this purpose, we applied NMF to the semantic analysis of a corpus of encyclopedia articles.

Some of the semantic features ($r = 200$, columns of W) discovered by NMF are shown in Fig. 4. In each semantic feature, the algorithm has grouped together semantically related words. Each article in the encyclopedia is represented by additively combining several of these features. For example, to represent the article about the 'Constitution of the United States', the semantic feature containing 'supreme' and 'court' and the one containing 'president' and 'congress' are coactivated.

In addition to grouping semantically related words together into semantic features, the algorithm uses context to differentiate between multiple meanings of the same word. For example, the word 'lead' appears with high frequency in two semantic features shown in Fig. 4: it occurs with 'metal', 'copper' and 'steel' in one, whereas it appears with 'person', 'rules' and 'law' in the other. This demonstrates that NMF can deal with the polysemy of 'lead' by disambiguating two of its meanings in the corpus of documents.

Although NMF is successful in learning facial parts and semantic topics, this success does not imply that the method can learn parts from any database, such as images of objects viewed from extremely different viewpoints, or highly articulated objects. Learning parts for these complex cases is likely to require fully hierarchical models with multiple levels of hidden variables, instead of the single level in NMF. Although non-negativity constraints may help such models to learn parts-based representations¹³, we do not claim that they are sufficient in themselves. Also, NMF does not learn anything about the 'syntactic' relationships between parts. NMF assumes that the hidden variables are non-negative, but makes no further assumptions about their statistical dependencies.

This is in contrast to independent components analysis (ICA), a variant of PCA that assumes that the hidden variables are statistically independent and non-gaussian^{16,12}. Applying ICA to the facial images to make the encodings independent results in basis images that are holistic. The independence assumption of ICA is ill-suited for learning parts-based representations because various

$$W_{ia} \leftarrow W_{ia} \sum_{\mu} \frac{V_{i\mu}}{(WH)_{i\mu}} H_{a\mu}$$

$$W_{ia} \leftarrow \frac{W_{ia}}{\sum_j W_{ja}}$$

$$H_{a\mu} \leftarrow H_{a\mu} \sum_i W_{ia} \frac{V_{i\mu}}{(WH)_{i\mu}}$$

Figure 2 Iterative algorithm for non-negative matrix factorization. Starting from non-negative initial conditions for W and H , iteration of these update rules for non-negative V finds an approximate factorization $V \approx WH$ by converging to a local maximum of the objective function given in equation (2). The fidelity of the approximation enters the updates through the quotient $V_{i\mu}/(WH)_{i\mu}$. Monotonic convergence can be proven using techniques similar to those used in proving the convergence of the EM algorithm^{22,23}. The update rules preserve the non-negativity of W and H and also constrain the columns of W to sum to unity. This sum constraint is a convenient way of eliminating the degeneracy associated with the invariance of WH under the transformation $W \rightarrow W\Lambda$, $H \rightarrow \Lambda^{-1}H$, where Λ is a diagonal matrix.

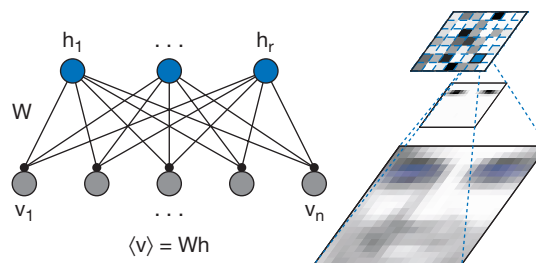


Figure 3 Probabilistic hidden variables model underlying non-negative matrix factorization. The model is diagrammed as a network depicting how the visible variables $v_1, \dots, v_n$ in the bottom layer of nodes are generated from the hidden variables $h_1, \dots, h_r$ in the top layer of nodes. According to the model, the visible variables v_i are generated from a probability distribution with mean $\sum_a W_{ia} h_a$. In the network diagram, the influence of h_a on v_i is represented by a connection with strength W_{ia} . In the application to facial images, the visible variables are the image pixels, whereas the hidden variables contain the parts-based encoding. For fixed a , the connection strengths $W_{1a}, \dots, W_{na}$ constitute a specific basis image (right middle) which is combined with other basis images to represent a whole facial image (right bottom).

parts are likely to occur together. This results in complex dependencies between the hidden variables that cannot be captured by algorithms that assume independence in the encodings. An alternative application of ICA is to transform the PCA basis images, to make the images rather than the encodings as statistically independent as possible¹⁸. This results in a basis that is non-global; however, in this representation all the basis images are used in cancelling combinations to represent an individual face, and thus the encodings are not sparse. In contrast, the NMF representation contains both a basis and encoding that are naturally sparse, in that many of the components are exactly equal to zero. Sparseness in both the basis and encodings is crucial for a parts-based representation.

The algorithm of Fig. 2 performs both learning and inference simultaneously. That is, it both learns a set of basis images and infers values for the hidden variables from the visible variables. Although the generative model of Fig. 3 is linear, the inference computation is nonlinear owing to the non-negativity constraints. The computation is similar to maximum likelihood reconstruction in emission tomography¹⁹, and deconvolution of blurred astronomical images^{20,21}.

According to the generative model of Fig. 3, visible variables are generated from hidden variables by a network containing excitatory connections. A neural network that infers the hidden from the visible variables requires the addition of inhibitory feedback connections. NMF learning is then implemented through plasticity in the synaptic connections. A full discussion of such a network is beyond the scope of this letter. Here we only point out the

consequence of the non-negativity constraints, which is that synapses are either excitatory or inhibitory, but do not change sign. Furthermore, the non-negativity of the hidden and visible variables corresponds to the physiological fact that the firing rates of neurons cannot be negative. We propose that the one-sided constraints on neural activity and synaptic strengths in the brain may be important for developing sparsely distributed, parts-based representations for perception. □

Methods

The facial images used in Fig. 1 consisted of frontal views hand-aligned in a 19×19 grid. For each image, the greyscale intensities were first linearly scaled so that the pixel mean and standard deviation were equal to 0.25, and then clipped to the range [0,1]. NMF was performed with the iterative algorithm described in Fig. 2, starting with random initial conditions for W and H . The algorithm was mostly converged after less than 50 iterations; the results shown are after 500 iterations, which took a few hours of computation time on a Pentium II computer. PCA was done by diagonalizing the matrix VV^T . The 49 eigenvectors with the largest eigenvalues are displayed. VQ was done via the k -means algorithm, starting from random initial conditions for W and H .

In the semantic analysis application of Fig. 4, the vocabulary was defined as the 15,276 most frequent words in the database of Grolier encyclopedia articles, after removal of the 430 most common words, such as 'the' and 'and'. Because most words appear in relatively few articles, the word count matrix V is extremely sparse, which speeds up the algorithm. The results shown are after the update rules of Fig. 2 were iterated 50 times starting from random initial conditions for W and H .

Received 24 May; accepted 6 August 1999.

- Palmer, S. E. Hierarchical structure in perceptual representation. *Cogn. Psychol.* **9**, 441–474 (1977).
- Wachsmuth, E., Oram, M. W. & Perrett, D. I. Recognition of objects and their component parts: responses of single units in the temporal cortex of the macaque. *Cereb. Cortex* **4**, 509–522 (1994).
- Logothetis, N. K. & Sheinberg, D. L. Visual object recognition. *Annu. Rev. Neurosci.* **19**, 577–621 (1996).
- Biederman, I. Recognition-by-components: a theory of human image understanding. *Psychol. Rev.* **94**, 115–147 (1987).
- Ullman, S. *High-Level Vision: Object Recognition and Visual Cognition* (MIT Press, Cambridge, MA, 1996).
- Turk, M. & Pentland, A. Eigenfaces for recognition. *J. Cogn. Neurosci.* **3**, 71–86 (1991).
- Field, D. J. What is the goal of sensory coding? *Neural Comput.* **6**, 559–601 (1994).
- Foldiak, P. & Young, M. Sparse coding in the primate cortex. *The Handbook of Brain Theory and Neural Networks* 895–898 (MIT Press, Cambridge, MA, 1995).
- Olshausen, B. A. & Field, D. J. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature* **381**, 607–609 (1996).
- Lee, D. D. & Seung, H. S. Unsupervised learning by convex and conic coding. *Adv. Neural Info. Proc. Syst.* **9**, 515–521 (1997).
- Paatero, P. Least squares formulation of robust non-negative factor analysis. *Chemometr. Intell. Lab. Syst.* **37**, 23–35 (1997).
- Nakayama, K. & Shimojo, S. Experiencing and perceiving visual surfaces. *Science* **257**, 1357–1363 (1992).
- Hinton, G. E., Dayan, P., Frey, B. J. & Neal, R. M. The "wake-sleep" algorithm for unsupervised neural networks. *Science* **268**, 1158–1161 (1995).
- Salton, G. & McGill, M. J. *Introduction to Modern Information Retrieval* (McGraw-Hill, New York, 1983).
- Landauer, T. K. & Dumais, S. T. The latent semantic analysis theory of knowledge. *Psychol. Rev.* **104**, 211–240 (1997).
- Jutten, C. & Herault, J. Blind separation of sources, part I: An adaptive algorithm based on neuromimetic architecture. *Signal Proc.* **24**, 1–10 (1991).
- Bell, A. J. & Sejnowski, T. J. An information maximization approach to blind separation and blind deconvolution. *Neural Comput.* **7**, 1129–1159 (1995).
- Bartlett, M. S., Lades, H. M. & Sejnowski, T. J. Independent component representations for face recognition. *Proc. SPIE* **3299**, 528–539 (1998).
- Shepp, L. A. & Vardi, Y. Maximum likelihood reconstruction for emission tomography. *IEEE Trans. Med. Imaging* **2**, 113–122 (1982).
- Richardson, W. H. Bayesian-based iterative method of image restoration. *J. Opt. Soc. Am.* **62**, 55–59 (1972).
- Lucy, L. B. An iterative technique for the rectification of observed distributions. *Astron. J.* **74**, 745–754 (1974).
- Dempster, A. P., Laird, N. M. & Rubin, D. B. Maximum likelihood from incomplete data via the EM algorithm. *J. Royal Stat. Soc.* **39**, 1–38 (1977).
- Saul, L. & Pereira, F. *Proceedings of the Second Conference on Empirical Methods in Natural Language Processing* (eds Cardie, C. & Weischedel, R.) 81–89 (Morgan Kaufmann, San Francisco, 1997).

Acknowledgements

We acknowledge the support of Bell Laboratories and MIT. C. Papageorgiou and T. Poggio provided us with the database of faces, and R. Sproat with the Grolier encyclopedia corpus. We thank L. Saul for convincing us of the advantages of EM-type algorithms. We have benefited from discussions with B. Anderson, K. Clarkson, R. Freund, L. Kaufman, E. Rietman, S. Roweis, N. Rubin, J. Tenenbaum, N. Tishby, M. Tsodyks, T. Tyson and M. Wright.

Correspondence and requests for materials should be addressed to H.S.S.

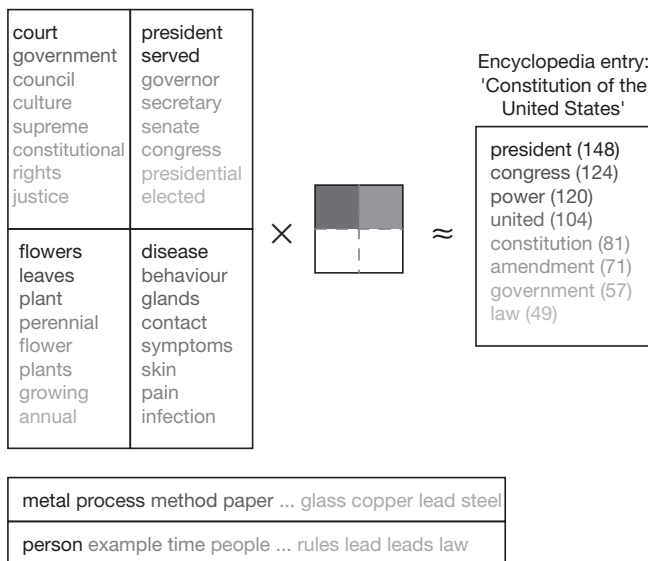


Figure 4 Non-negative matrix factorization (NMF) discovers semantic features of $m = 30,991$ articles from the Grolier encyclopedia. For each word in a vocabulary of size $n = 15,276$, the number of occurrences was counted in each article and used to form the $15,276 \times 30,991$ matrix V . Each column of V contained the word counts for a particular article, whereas each row of V contained the counts of a particular word in different articles. The matrix was approximately factorized into the form WH using the algorithm described in Fig. 2. Upper left, four of the $r = 200$ semantic features (columns of W). As they are very high-dimensional vectors, each semantic feature is represented by a list of the eight words with highest frequency in that feature. The darkness of the text indicates the relative frequency of each word within a feature. Right, the eight most frequent words and their counts in the encyclopedia entry on the 'Constitution of the United States'. This word count vector was approximated by a superposition that gave high weight to the upper two semantic features, and none to the lower two, as shown by the four shaded squares in the middle indicating the activities of H . The bottom of the figure exhibits the two semantic features containing 'lead' with high frequencies. Judging from the other words in the features, two different meanings of 'lead' are differentiated by NMF.

Distributed synaptic modification in neural networks induced by patterned stimulation

Guo-qiang Bi & Mu-ming Poo

Department of Biology, University of California at San Diego, La Jolla, California 92093-0357, USA

Activity-dependent changes in synaptic efficacy or connectivity are critical for the development¹, signal processing² and learning and memory functions^{3–6} of the nervous system. Repetitive correlated spiking of pre- and postsynaptic neurons can induce a persistent increase or decrease in synaptic strength, depending on the timing of the pre- and postsynaptic excitation^{7–13}. Previous studies on such synaptic modifications have focused on synapses made by the stimulated neuron. Here we examine, in networks of cultured hippocampal neurons, whether and how localized stimulation can modify synapses that are remote from the stimulated neuron. We found that repetitive paired-pulse stimulation of a single neuron for brief periods induces persistent strengthening or weakening of specific polysynaptic pathways in a manner that depends on the interpulse interval. These changes can be accounted for by correlated pre- and postsynaptic excitation at distant synaptic sites, resulting from different transmission delays along separate pathways. Thus, through such a 'delay-line' mechanism, temporal information coded in the timing of individual spikes^{14–17} can be converted into and stored as spatially distributed patterns of persistent synaptic modifications in a neural network.

Cultures of dissociated rat hippocampal neurons were used to study functional networks consisting of both glutamate- and γ -aminobutyric acid (GABA)-mediated synapses¹³. Perforated whole-cell recordings of synaptic currents were made from 2–3 neurons in networks of about 10–20 neurons. The evoked postsynaptic currents (PSCs) elicited by a brief depolarizing stimulus (+100 mV,

1 ms) applied to another neuron, or to the recorded neuron itself, usually exhibited complex but reproducible patterns with distinct components (Fig. 1a). Each PSC component corresponds to a polysynaptic pathway with a specific transmission delay, which is due to synaptic delays and the time required for the initiation and conduction of action potentials, estimated to be in the range of 4–10 ms per synapse (see Methods). Under low-frequency test stimulation, the average probability of occurrence (P) for each PSC component remained relatively constant throughout the recording period (Fig. 1b). The PSC profile thus provides a reliable means for monitoring the status of synaptic efficacy along multiple polysynaptic pathways in the network.

To examine the effect of repetitive local stimulation on remote synapses in the polysynaptic pathway, we applied a train of paired-pulse stimuli (PPS) with a defined interpulse interval (IPI) to a single neuron in the network (voltage clamp $V_c = -70$ mV). Paired pulses were chosen because of their simple temporal structure. After 20 paired pulses at 1 Hz, we frequently observed changes in the PSC profile in either the stimulated neuron (Fig. 2a, c–f) or a different neuron in the network (Fig. 2b). These changes included the appearance of new components (Fig. 2a, d) and the disappearance or change in P of pre-existing components (Fig. 2b–d), reflecting remodelling of polysynaptic pathways in the network. Such pathway remodelling was persistent, lasting for as long as stable recording was made (up to 1.5 h), indicating that long-term changes had occurred in specific pathways between the stimulated and recorded neurons. In Fig. 2d, repetitive PPS induced the appearance of a new component 3, and an increase and decrease in P for components 2 and 1, respectively; thus, the same pattern of stimuli can induce opposite changes along different pathways. In most cases, changes in the PSC profile involved alterations in P for particular components without significant changes in the amplitude. Thus, the patterned stimulation had apparently resulted in changes at remote synaptic connections, thereby changing the probability of successful transmission along different pathways leading to the recorded neuron. Such distributed changes are consistent with the notion of distributed representation and storage of information in neural networks^{3,4}.

When multiple episodes of repetitive PPS (20 pairs, 1 Hz) were applied to the same neuron, pathway remodelling depended on the

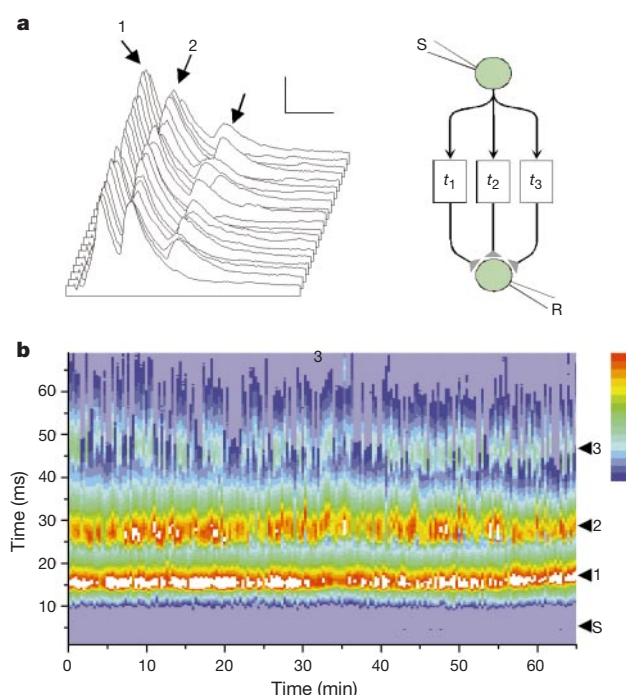


Figure 1 Polysynaptic pathways in a network of cultured hippocampal neurons. **a**, Traces depict 20 consecutive postsynaptic currents (PSCs) (inward currents are shown upwards) recorded from a hippocampal neuron in response to test stimulation (0.05 Hz) of a nearby neuron (see Methods). The diagram on the right depicts three hypothetical polysynaptic pathways leading to the recorded neuron with transmission delays of t_1 , t_2 and t_3 , respectively, corresponding to the onset latencies of the three distinct PSC components. Scales: 20 ms, 500 pA. **b**, The stability of polysynaptic pathways. Shown are 195 consecutive PSCs, each triggered by a test stimulus and represented by a single vertical line, with current magnitude coded by colour. Colour code: linear scale 150–700 pA; white and background colours represent above- and below-scale values, respectively. Arrowheads indicate the time of stimulation (S) and peaks of identified PSC components (1–3).

IPI of the paired pulses. In the experiment shown in Fig. 2e, no persistent change in PSCs was induced by paired pulses with 60 ms IPI, but paired pulses of 40 ms IPI induced two new PSC components with a relatively high P , which persisted for more than 1 h. In another experiment (Fig. 2f), the first episode of stimulation (IPI, 100 ms) had no significant effect, but the second episode with 50 ms IPI resulted in an increased P for two components, one of which was reduced by further stimulation with a third episode (IPI 20 ms). In a third experiment (see Supplementary Information), four separate episodes were applied to the network to show further the specificity of the IPI in inducing pathway remodelling. Therefore, pathway remodelling is highly dependent on the precise temporal pattern of the repetitive localized stimulation.

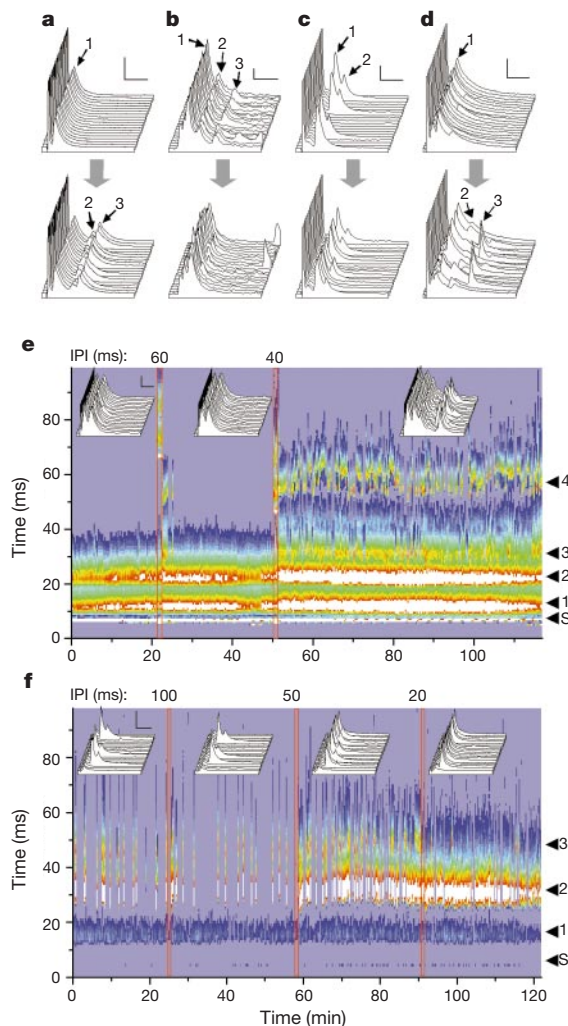


Figure 2 Pathway remodelling induced by repetitive paired-pulse stimulation (PPS). **a–d**, Four experiments showing different types of change in the PSC profile. For each experiment, consecutive PSCs elicited by test stimuli before (top) and 10 min after (bottom) PPS (20 pairs, 1 Hz) with interpulse interval (IPI) of 50 (**a**), 30 (**b**), 50 (**c**) and 100 ms (**d**) are shown. PSCs were recorded either in the stimulated neuron (**a**, **c**, **d**: note the large stimulus artifact) or in a nearby neuron (**b**). Scales: 20 ms, 1 nA (**a**, **b**); 20 ms, 500 pA (**c**, **d**). **e**, **f**, Changes in the PSC profile depend on the IPI of the paired pulses. **e**, Two episodes of PPS (marked by vertical red lines) applied to the recorded neuron; insets, 20 consecutive PSCs at 10–16, 40–46, and 80–86 min. Scales: 20 ms, 500 pA. Colour code as in Fig. 1b: 0.2–1.2 nA, linear scale. **f**, PSCs elicited by stimulating a different neuron in the network (test frequency 0.07 Hz); insets, samples of PSCs at 10–14, 40–44, 80–84 and 110–114 min. Scale: 20 ms, 400 pA, colour code: 50–400 pA, linear scale.

Repetitive correlated pre- and postsynaptic excitation can induce long-term synaptic modification, and the direction of synaptic changes depends on the temporal order of pre- and postsynaptic activation^{7–13}. In the same hippocampal cultures¹³, postsynaptic spiking within 20 ms after presynaptic activation (positively correlated

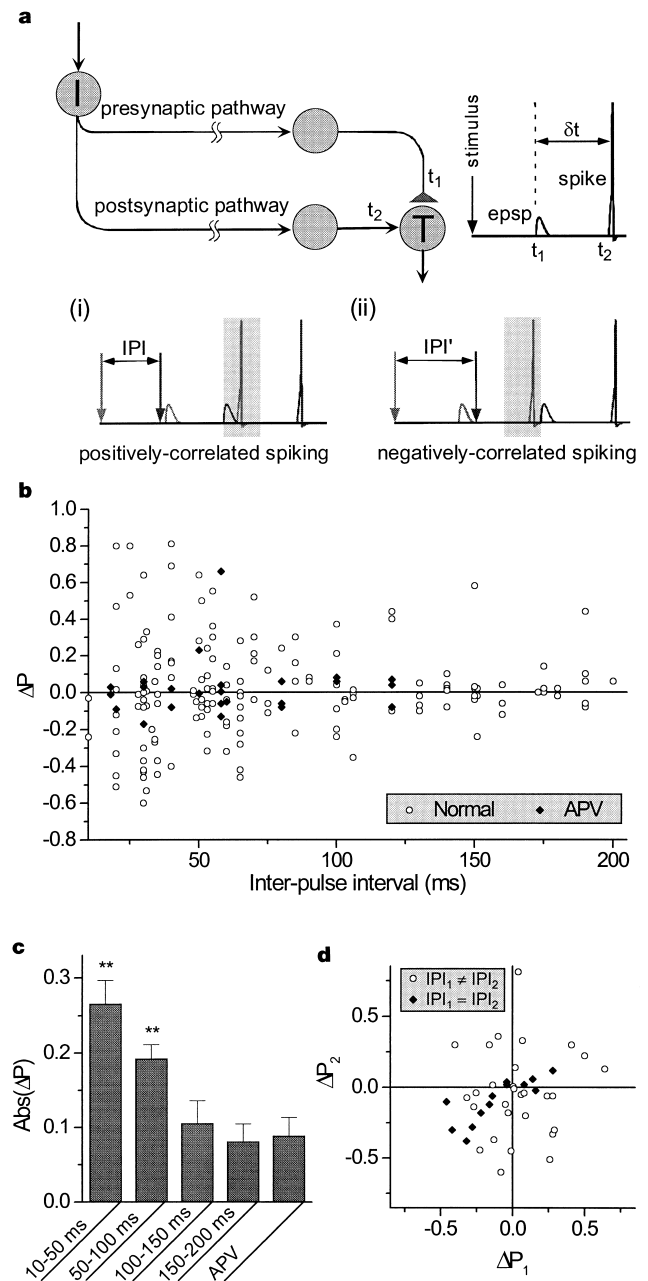


Figure 3 Pathway remodelling induced by correlated excitation through transmission delay. **a**, Correlated pre- and postsynaptic excitation at a 'remote' synapse (triangle) by stimulation transmitted along two converging pathways (see text). The 'postsynaptic' synapse (with respect to the remote synapse) can trigger spikes at the target neuron. Shaded areas indicate effective time windows (~ 20 ms) for synaptic modifications¹³. **b**, **c**, Summary of pathway remodelling induced by PPS (20 pairs, 1 Hz) of various IPIs. Each point represents the result for one identified PSC component from a single experiment without (Normal) or with 50 μ M D-AP-5 (APV). Columns in **c** represent averages of absolute values of ΔP (y axis) from experiments with different ranges of IPIs (x axis). Two asterisks, significantly different from the APV group ($P < 0.005$, t -test). **d**, Correlation between changes in probability of occurrence (ΔP_1 and ΔP_2) after two consecutive episodes of PPS with the same or different IPIs. Correlation coefficients: $\rho = 0.058$ for $IPI_1 \neq IPI_2$ ($P > 0.25$, t -test); $\rho = 0.81$ for $IPI_1 = IPI_2$ ($P < 0.001$, t -test).

spiking) leads to long-term potentiation (LTP), whereas post-synaptic spiking within 20 ms before presynaptic activation (negatively correlated spiking) leads to long-term depression (LTD). Thus, paired-pulse facilitation or temporal integration of the evoked postsynaptic potentials (EPSPs) at the remote synapses might generate positively correlated spiking at these synapses, resulting in pathway remodelling by homosynaptic potentiation. However, the prevailing paired-pulse depression in these cultures (see Supplementary Information) makes such a process unlikely. An

alternative mechanism that might generate correlated spiking at remote synapses is based on different transmission delays along converging pathways. As shown in Fig. 3a, localized stimulation of an input neuron (I) results in activation of two separate pathways that converge on a target neuron (T) with a difference in the transmission delay of $t_2 - t_1$ (or ' δt '). For the case of $t_2 > t_1$, repetitive PPS with IPI $< \delta t$ may result in positively correlated spiking, and hence LTP at the remote synapse made by the presynaptic pathway onto neuron T; whereas stimulation with

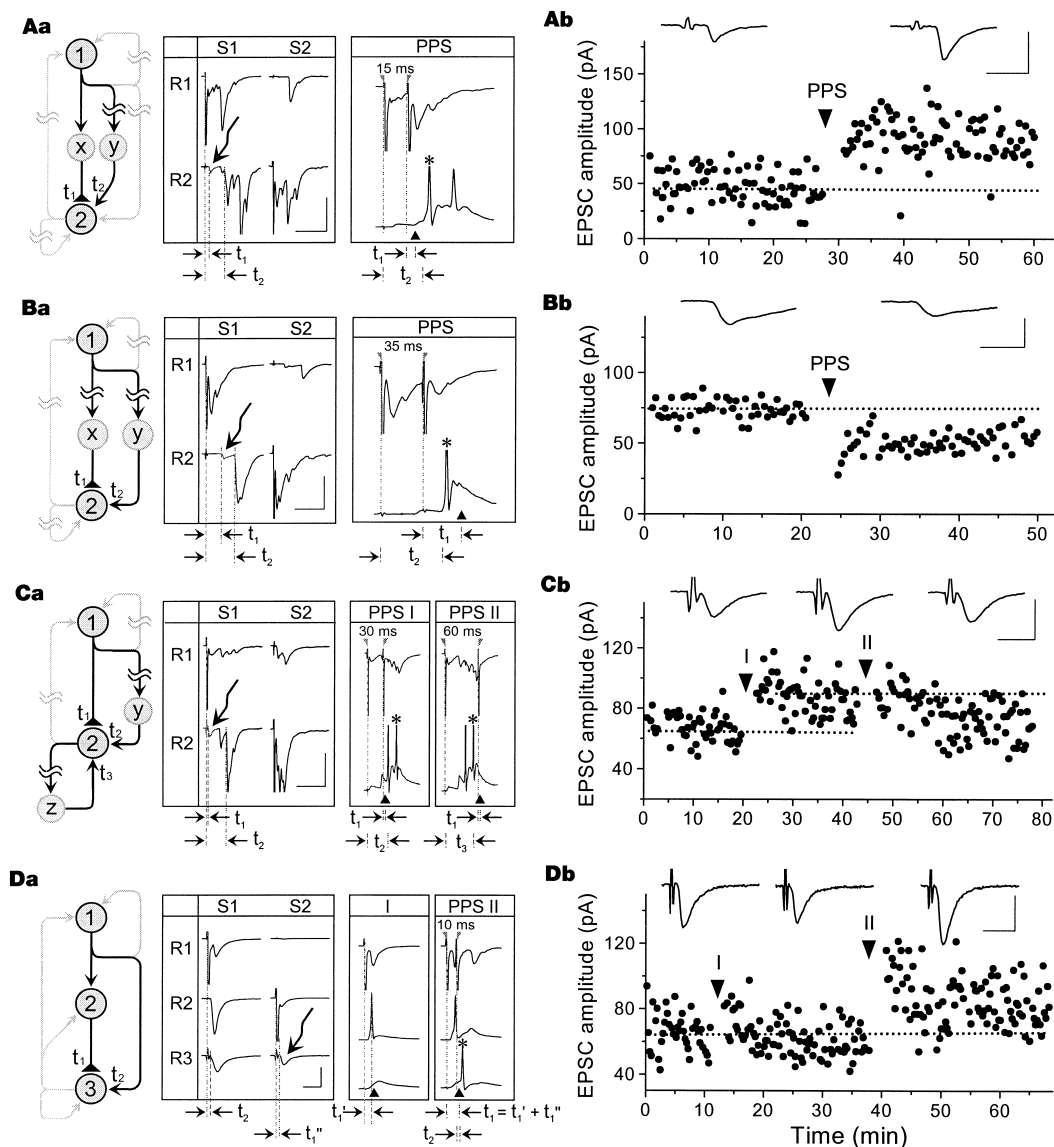


Figure 4 Modifications of remote excitatory synapses by correlated spiking. **Aa**, Synaptic connections between two recorded neurons (left panel). Dark lines, two convergent pathways of interest (transmission delays: $t_1 = 6$ ms; $t_2 = 27$ ms); grey lines, other detected connections; cells x, y, unrecorded neurons upstream of cell 2; triangle, the 'remote' synapse ($E_{x \rightarrow 2}$) under examination; $\approx$ denotes polysynaptic connections. PSCs were recorded from cells 1 and 2 (R1 and R2) when they were stimulated (S1 and S2) separately (middle panel). Arrow points to the PSC component corresponding to $E_{x \rightarrow 2}$. Synaptic responses recorded from cells 1 (top) and 2 (bottom) during repetitive PPS (1 Hz, 50 s) of cell 1 while cell 2 was current-clamped (right panel). Note expanded timescale. Scales: 40 ms, 500 pA for EPSCs in S1, S2; 20 ms, 50 mV for EPSPs in PPS. **Ab**, LTP at $E_{x \rightarrow 2}$ resulting from positively correlated spiking during PPS for the circuit shown in **Aa**. Test stimuli (0.05 Hz) were applied to cell 1. Insets, sample EPSCs (mean of 10 consecutive events) 10 min before and 20 min after PPS. Scales: 10 ms, 100 pA. **B**, LTD

at $E_{x \rightarrow 2}$ ($t_1 = 28$ ms) by negatively correlated spiking through $E_{y \rightarrow 2}$ ($t_2 = 51$ ms). Scales: 50 ms, 500 pA for EPSCs, 25 ms, 50 mV for EPSPs (**Ba**), 10 ms, 100 pA (**Bb**). **C**, LTP and LTD at $E_{1 \rightarrow 2}$ (monosynaptic, $t_1 = 3$ ms) by correlated spiking through $E_{y \rightarrow 2}$ and $E_{z \rightarrow 2}$ ($t_2 = 36$ ms; $t_3 = 50$ ms). During the two episodes of PPS, recurrent excitation (through $E_{z \rightarrow 2}$) caused a second spike at cell 2. Scales: 50 ms, 500 pA for EPSCs, 50 ms, 50 mV for EPSPs (**Ca**); 10 ms, 100 pA (**Cb**). **D**, A serially connected triplet with all neurons recorded. Synapse $E_{1 \rightarrow 2}$ was suprathreshold. Transmission delay of the $1 \rightarrow 2 \rightarrow 3$ pathway ($t_1 = 12$ ms) was due to the latency of synaptic delay and spike initiation at cell 2 ($t'_1 = 8$ ms), and the synaptic delay at $E_{2 \rightarrow 3}$ ($t'_2 = 4$ ms). Two episodes of repetitive stimulation (I, 50 single pulses; II, 50 paired pulses; IPI 10 ms, both at 1 Hz) were applied to cell 1 while cell 2 and 3 were under current clamp. Scales: 20 ms, 200 pA for EPSCs, 10 ms, 50 mV for EPSPs (**Da**); 20 ms, 50 pA (**Db**).

$IPI > \delta t$ may result in negatively correlated spiking, and hence LTD at that synapse. Modification of this remote synapse may change the firing probability of neuron T, leading to a change in the corresponding PSC components in downstream neurons. Such a heterosynaptic mechanism is sensitive to the IPI of the paired-pulse stimuli and may thus account for the pathway remodelling described above.

The effects of paired pulses with IPIs of 10–200 ms were examined. In each experiment, we started with an IPI within this range and tested 2–4 episodes of 20 paired pulses (at 1 Hz) of progressively decreasing or increasing IPIs. Before and after each episode, PSCs were monitored by low-frequency test stimuli (at 0.05 Hz) for 20–40 min. When changes in the probability of occurrence (ΔP) of identified PSC components were plotted against the IPIs of the stimuli (Fig. 3b), we found that the changes decreased with increasing IPI, with IPIs below 100 ms being most effective in inducing pathway remodelling (Fig. 3b, c). This is consistent with a decreasing probability of forming converging pathways with longer pathlength differences that are responsible for generating correlated spiking by PPS in the cultured networks.

We further compared the results of applying two consecutive episodes of paired pulses of the same IPI with those of different IPIs. As shown in Fig. 3d, a strong correlation was observed between changes in P (ΔP_1 and ΔP_2) after the first and the second episode of the same IPIs. In contrast, two episodes of IPIs differing by 10–40 ms yielded results that showed little correlation, indicating the network's capability of discriminating a temporal difference of ~ 10 ms. These results support the heterosynaptic mechanism shown in Fig. 3a.

Long-term modification of many central synapses depends on the activation of *N*-methyl-D-aspartate (NMDA) subtypes of glutamate receptors^{5,18,19}. As shown in Fig. 3b, significantly fewer changes in the PSC profile were induced by PPS in the presence of D-2-amino-5-phosphonopentanoic acid (D-AP5; 50 μ M), a specific blocker of NMDA receptors. The D-AP5 treatment did not significantly affect the network activity, as shown by comparison of the PSC profile before and after the treatment (data not shown), but the reduction in pathway remodelling in the presence of D-AP5 is consistent with the involvement of NMDA receptor-dependent LTP and LTD¹³. We did observe two cases of significant changes ($>20\%$) in P of PSC components in the presence of D-AP5. This may reflect other NMDA-receptor-independent mechanisms, such as modifications of inhibitory GABA-mediated synapses^{19,20} or activity-induced changes in cell excitability²¹.

To verify directly that remote synaptic changes can be induced by correlated excitation through converging pathways, we examined modifications of identified PSC components by PPS when the postsynaptic cell was under current clamp. As shown in Fig. 4A–C, the relevant transmission delays (t_1 , t_2 , t_3) along identified pathways from cell 1 to cell 2 (equivalent to neuron T in Fig. 3a) can be precisely determined from the recorded excitatory synaptic inputs (designated as ' $E_{x \rightarrow 2}$ ', ' $E_{y \rightarrow 2}$ ' and ' $E_{z \rightarrow 2}$ '). In these experiments, during repetitive PPS, the postsynaptic spike (marked by an asterisk) due to a suprathreshold input (' $E_{y \rightarrow 2}$ ' or ' $E_{z \rightarrow 2}$ ') triggered by the first stimulus was positively or negatively correlated with the EPSP (onset marked by arrowhead) of the 'remote' synaptic input ($E_{x \rightarrow 2}$ or $E_{1 \rightarrow 2}$) triggered by the second stimulus, resulting in LTP or LTD, respectively. Note that the EPSP triggered by the first stimulus fell outside the 20 ms window¹³ for synaptic modification by correlated spiking. The specificity of IPIs for modifying the identified synapses is consistent with the heterosynaptic mechanism shown in Fig. 3a. Experiments were also performed on simpler networks using simultaneous whole-cell recordings from three neurons. In the case shown in Fig. 4D, the serially connected triplet has two converging pathways onto neuron 3. Repetitive single-pulse stimulation of cell 1 had no significant effect on synapse $E_{2 \rightarrow 3}$, although synapse $E_{1 \rightarrow 2}$ was potentiated as expected (data not

shown). However, stimulation of neuron 1 with paired pulses (IPI 10 ms) caused spiking at cell 3 due to summation of EPSPs from $E_{1 \rightarrow 3}$ and $E_{2 \rightarrow 3}$, resulting in LTP at $E_{2 \rightarrow 3}$. Together, these experiments confirm that repetitive local stimulation of a neuron can induce LTP and LTD at remote polysynaptic sites depending on the timing of correlated pre- and postsynaptic excitation, which in turn is determined by differential transmission delays along converging pathways.

Tetanic stimulation of afferent fibres to hippocampal CA3 neurons or perforant path (PP) fibres results in long-term activation of latent polysynaptic pathways²⁰ or potentiation of population spikes at existing polysynaptic pathways²², respectively. The underlying mechanisms for these phenomena appeared to be a change in inhibitory control²⁰ or synaptic coupling of tetanizing trains through direct PP–CA3/CA1 synapses with asynchronous polysynaptic volleys occurring in the intrahippocampal circuitry²². Our results from dissociated neuronal cultures provide a direct demonstration of a generic property of neural networks: temporal correlation through transmission delays can be used for selective pathway modifications in accordance with the precise temporal structure of the input stimuli. Such a strategy of temporal-to-spatial conversion, different from that using paired-pulse facilitation and slow inhibition for selective activation in the hippocampal circuit²³, is reminiscent of using delay lines for parallel processing of temporally structured information in the auditory system^{24,25} and in neural network models^{16,26,27}. Here, polysynaptic pathways may be regarded as delay lines with a temporal resolution in the order of milliseconds and a range of delay times proportional to the size of the network. In intact nervous systems, synaptic input from a single neuron is usually not sufficient to trigger postsynaptic spiking. Thus, single neurons serially connected may not form functional polysynaptic pathways. However, the firing of a cultured neuron hyperinnervating others can be considered to be analogous to the *in vivo* situation of synchronous firing of a group of neurons with common targets^{15,28}. Polysynaptic transmission pathways in culture therefore simulate *in vivo* pathways which mediate information flow by correlated firing^{15,28}. If spike timing encodes neural information^{14–17}, the delay line architecture combined with spike-timing-based synaptic modifications provide a network mechanism to convert and store temporal information into spatially distributed patterns of synaptic modifications. □

Methods

Cell culture

Low-density cultures of dissociated embryonic rat hippocampal neurons were prepared as described^{13,29}. Hippocampi were removed from E18–20 embryonic rats and treated with trypsin for 15 min at 37 °C, followed by washing and gentle trituration. The dissociated cells were plated on poly-L-lysine-coated glass coverslips in 35-mm petri dishes at densities of 30,000–90,000 cells per dish. The plating medium was Dulbecco's Minimum Essential Medium (DMEM, BioWhittaker) supplemented with 10% heat-inactivated fetal bovine serum (Hyclone), 10% Ham's F12 with glutamine (BioWhittaker) and 50 U ml^{−1} penicillin-streptomycin (Sigma). Twenty-four hours after plating, one-third of the culture medium was replaced by the same medium supplemented with 20 mM KCl. Cells were used for electrophysiological studies after 10–15 days in culture, when functional neuronal networks had been established. A typical network selected consisted of about 10–20 neurons on a patch of glial cell monolayer of ~ 1 –2 mm² in size. Larger networks in older cultures usually exhibit unpredictable bursts of spontaneous activity and complex PSC profiles, and were thus avoided.

Electrophysiology

Whole-cell perforated patch recordings³⁰ from 2–3 hippocampal neurons were performed simultaneously, using amphotericin B (150 μ g ml^{−1}, Calbiochem) for perforation. The micropipettes were made from borosilicate glass capillaries (VWR), with a resistance in the range of 1.5–3 M Ω . The internal solution contained the following (in mM): K-gluconate 136.5, KCl 17.5, NaCl 9, MgCl₂ 1, HEPES 10, EGTA 0.2, pH 7.2. The external bath solution was a HEPES-buffered saline (HBS) containing the following (in mM): NaCl 150, KCl 3, CaCl₂ 3, MgCl₂ 2, HEPES 10, glucose 5, pH 7.3. For blocking NMDA receptors, 50 μ M D-AP5 (RPI) was added to HBS. The culture was constantly perfused with fresh external solution at ~ 1 ml min^{−1} at room temperature throughout the recording period.

The neurons were visualized by phase-contrast microscopy with an inverted microscope (Nikon Diaphot) while recording (in voltage clamp unless otherwise stated ($V_c = -70$ mV) and stimulation (1 ms, +100 mV step depolarization in voltage clamp) was performed using patch-clamp amplifiers (Axopatch 200B, Axon) interfaced with a PC. Signals filtered at 5 kHz using amplifier circuitry were sampled at 10 kHz and analysed using Axoscope software (Axon). Series resistance (10–30 M Ω) was always compensated at 80% (lag 100 μ s). For assaying synaptic connectivity, each neuron was stimulated at a low frequency (0.03–0.06 Hz), and the responses from the other neurons as well as autaptic responses in the stimulated neuron itself were recorded (see Fig. 4). Monosynaptic currents had onset latencies <4 ms (ref. 29). Polysynaptic currents had onset latencies ≥ 6 ms and often exhibited multiple components, with frequent failure for some PSC components to occur during the test stimulation. P of an identified component is measured based on its response to 50–100 consecutive test stimuli (0.05 Hz unless otherwise indicated). For polysynaptic pathways in these cultures, each additional synapse usually introduced a delay of 4–10 ms (2–5 ms for synaptic delay and action-potential conduction and 2–5 ms delay for the initiation of an action potential). Neurons in these cultures were either glutamate-mediating or GABA-mediating in nature and could be identified based on the time course, reversal potential and pharmacology of their evoked synaptic currents (EPSCs and IPSCs, respectively)^{13,29}. EPSCs were blocked by 10 μ M 6-cyano-7-nitroquinoxaline-2,3-dione (CNQX, RBI) whereas IPSCs were blocked by 10 μ M bicuculline (RBI). The IPSCs had distinctly longer decay time and more negative reversal potentials (around –50 mV) than EPSCs. In a typical culture, we estimated that less than 20% of the neurons were GABA-mediated.

Received 29 June; accepted 20 August 1999.

- Katz, L. C. & Shatz, C. J. Synaptic activity and the construction of cortical circuits. *Science* **274**, 1133–1138 (1996).
- Abbott, L. F., Varela, J. A., Sen, K. & Nelson, S. B. Synaptic depression and cortical gain control. *Science* **275**, 220–224 (1997).
- Squire, L. R. *Memory and Brain* (Oxford Univ. Press, New York, 1987).
- Churchland, P. S. & Sejnowski, T. J. *The Computational Brain* (The MIT Press, Cambridge, MA, 1992).
- Bliss, T. V. & Collingridge, G. L. A synaptic model of memory: long-term potentiation in the hippocampus. *Nature* **361**, 31–39 (1993).
- Goda, Y. & Stevens, C. F. Synaptic plasticity: the basis of particular types of learning. *Curr. Biol.* **6**, 375–378 (1996).
- Levy, W. B. & Steward, O. Temporal contiguity requirements for long-term associative potentiation/depression in the hippocampus. *Neuroscience* **8**, 791–797 (1983).
- Markram, H., Lübke, J., Frotscher, M. & Sakmann, B. Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. *Science* **275**, 213–215 (1997).
- Magee, J. C. & Johnston, D. A synaptically controlled, associative signal for Hebbian plasticity in hippocampal neurons. *Science* **275**, 209–213 (1997).
- Bell, C. C., Han, V. Z., Sugawara, Y. & Grant, K. Synaptic plasticity in a cerebellum-like structure depends on temporal order. *Nature* **387**, 278–281 (1997).
- Debanne, D., Gähwiler, B. H. & Thompson, S. M. Long-term synaptic plasticity between pairs of individual CA3 pyramidal cells in rat hippocampal slice cultures. *J. Physiol. (Lond.)* **507**, 237–247 (1998).
- Zhang, L. I., Tao, H. W., Holt, C. E., Harris, W. A. & Poo, M.-m. A critical window for cooperation and competition among developing retinotectal synapses. *Nature* **395**, 37–44 (1998).
- Bi, G.-q. & Poo, M.-m. Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type. *J. Neurosci.* **18**, 10464–10472 (1998).
- Hopfield, J. J. Pattern recognition computation using action potential timing for stimulus representation. *Nature* **376**, 33–36 (1995).
- Singer, W. & Gray, C. M. Visual feature integration and the temporal correlation hypothesis. *Annu. Rev. Neurosci.* **18**, 555–586 (1995).
- Gerstner, W., Kempter, R., van Hemmen, J. L. & Wagner, H. A neuronal learning rule for sub-millisecond temporal coding. *Nature* **383**, 76–81 (1996).
- Strong, S. P., Koberle, R., De Ruyter Van Steveninck, R. R. & Bialek, W. Entropy and information in neural spike trains. *Phys. Rev. Lett.* **80**, 197–200 (1998).
- Malenka, R. C. & Nicoll, R. A. NMDA-receptor-dependent synaptic plasticity: multiple forms and mechanisms. *Trends Neurosci.* **16**, 521–527 (1993).
- Linden, D. J. & Connor, J. A. Long-term synaptic depression. *Annu. Rev. Neurosci.* **18**, 319–357 (1995).
- Miles, R. & Wong, R. K. Latent synaptic pathways revealed after tetanic stimulation in the hippocampus. *Nature* **329**, 724–726 (1987).
- Turrigiano, G., Abbott, L. F. & Marder, E. Activity-dependent changes in the intrinsic properties of cultured neurons. *Science* **264**, 974–977 (1994).
- Buzsáki, G. Polysynaptic long-term potentiation: a physiological role of the perforant path–CA3/CA1 pyramidal cell synapse. *Brain Res.* **455**, 192–195 (1988).
- Buonomano, D. V., Hickmott, P. W. & Merzenich, M. M. Context-sensitive synaptic plasticity and temporal-to-spatial transformations in hippocampal slices. *Proc. Natl Acad. Sci. USA* **94**, 10403–10408 (1997).
- Carr, C. E. & Konishi, M. A circuit for detection of interaural time differences in the brain stem of the barn owl. *J. Neurosci.* **10**, 3227–3246 (1990).
- Kristan, W. B. J. He's got rhythm: single neurons signal timing on a scale of seconds. *Nature Neurosci.* **1**, 643–645 (1998).
- Tank, D. W. & Hopfield, J. J. Neural computation by concentrating information in time. *Proc. Natl Acad. Sci. USA* **84**, 1896–1900 (1987).
- Moore, J. W., Choi, J.-S. & Brunzell, D. H. In *Timing of Behaviour* (eds Risenbaum, D. A. & Collyer, C. E.) 3–34 (MIT Press, Cambridge, MA, 1998).
- Abeles, M. *Corticonics* (Cambridge Univ. Press, Cambridge, 1991).
- Fitzsimonds, R. M., Song, H. J. & Poo, M.-m. Propagation of activity-dependent synaptic depression in simple neural networks. *Nature* **388**, 439–448 (1997).
- Rae, J., Cooper, K., Gates, P. & Watsky, M. Low access resistance perforated patch recordings using amphotericin B. *J. Neurosci. Methods* **37**, 15–26 (1991).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

Acknowledgements

We thank X. Wang for culture preparations and B. Berninger, W. Kristan, L. Zhang, A. Schinder and S. Andersen for helpful discussions and comments on the manuscript. Supported by grants from NIH (M.P.) and a President's Postdoctoral Fellowship from the University of California and a training grant from NIH (G.B.).

Correspondence and requests for materials should be addressed to G.B. (e-mail: gbi@ucsd.edu).

Benzodiazepine actions mediated by specific γ -aminobutyric acid_A receptor subtypes

Uwe Rudolph*, Florence Crestani*, Dietmar Benke*, Ina Brünig*, Jack A. Benson*, Jean-Marc Fritschy*, James R. Martin†, Horst Bluethmann† & Hanns Möhler*

* Institute of Pharmacology, University of Zürich, and Swiss Institute of Technology (ETH), Winterthurestrasse 190, CH-8057 Zürich, Switzerland

† Pharma Division, Preclinical Research, F. Hoffmann-La Roche Ltd, CH-4002 Basel, Switzerland

GABA_A (γ -aminobutyric acid_A) receptors are molecular substrates for the regulation of vigilance, anxiety, muscle tension, epileptogenic activity and memory functions, which is evident from the spectrum of actions elicited by clinically effective drugs acting at their modulatory benzodiazepine-binding site. Here we show, by introducing a histidine-to-arginine point mutation at position 101 of the murine $\alpha 1$ -subunit gene, that $\alpha 1$ -type GABA_A receptors, which are mainly expressed in cortical areas and thalamus¹, are rendered insensitive to allosteric modulation by benzodiazepine-site ligands, whilst regulation by the physiological neurotransmitter γ -aminobutyric acid is preserved. $\alpha 1$ (H101R) mice failed to show the sedative, amnesic and partly the anticonvulsant action of diazepam. In contrast, the anxiolytic-like, myorelaxant, motor-impairing and ethanol-potentiating effects were fully retained, and are attributed to the nonmutated GABA_A receptors found in the limbic system ($\alpha 2$, $\alpha 5$), in monoaminergic neurons ($\alpha 3$) and in motoneurons ($\alpha 2$, $\alpha 5$)¹. Thus, benzodiazepine-induced behavioural responses are mediated by specific GABA_A receptor subtypes in distinct neuronal circuits, which is of interest for drug design.

Fast synaptic inhibition in the mammalian brain is largely mediated by the activation of GABA_A receptors, which are heteromeric GABA-gated chloride channels². Their opening frequency is enhanced by agonists of the benzodiazepine site, which is the basis of their therapeutic effectiveness but also of their undesired side effects. The classical benzodiazepines such as diazepam interact indiscriminately with all benzodiazepine-sensitive GABA_A receptor subtypes ($\alpha 1$, $\alpha 2$, $\alpha 3$ and $\alpha 5$) with comparable affinities^{3,4}. These receptors have a histidine at a conserved position ($\alpha 1$ -H101, $\alpha 2$ -H101, $\alpha 3$ -H126 and $\alpha 5$ -H105). In contrast, the benzodiazepine-insensitive receptor subtypes in the brain ($\alpha 4$ and $\alpha 6$) have an arginine in the corresponding position.

Recombinant diazepam-sensitive receptors can be rendered diazepam-insensitive by replacing this histidine by arginine without altering the GABA sensitivity, as shown for the $\alpha 1$ subunit^{5,6} and the $\alpha 2$, $\alpha 3$ and $\alpha 5$ subunits⁷. The pharmacological significance of the predominant GABA_A receptor subtype in the brain, which contains

the $\alpha 1$ subunit^{1,8}, was evaluated by introducing the $\alpha 1$ (H101R) point mutation into the germline of mice by gene targeting ('knock-in'). A replacement vector RV-10, which contained the desired point mutation in exon 4 and a *loxP*-flanked neomycin resistance marker in intron 4, was electroporated into E14 embryonic stem (ES) cells. Correctly targeted ES cells containing the point mutation and the neo marker (Mutant allele 1, Fig. 1a) were injected into blastocysts. Mice carrying this mutant allele were bred to Ella-cre mice⁹. The Ella-cre transgene efficiently eliminated the *loxP*-flanked neomycin resistance cassette from the germ line⁹, thus generating Mutant allele 2 (Fig. 1a), which was bred on different backgrounds and used in all subsequent experiments. The Ella-cre transgene was bred out. The mutant alleles were confirmed by Southern blotting (Fig. 1b) and sequence analysis (Fig. 1c).

$\alpha 1$ (H101R) mice showed no overt distinctive phenotype and bred normally. Immunoblotting confirmed that the mutant $\alpha 1$ subunit and the other major GABA_A receptor subunits ($\alpha 2$, $\alpha 3$, $\alpha 5$, $\beta 2/3$ and $\gamma 2$) were expressed in the $\alpha 1$ (H101R) mice at normal levels (Fig. 2a). The immunohistochemical distribution of these subunits in $\alpha 1$ (H101R) mice was indistinguishable from that of the

wild type (see also Fig. 2b). To verify the diazepam-insensitivity of GABA_A $\alpha 1$ (H101R) receptors in the central nervous system, GABA_A receptors were immunoprecipitated from detergent extracts of whole brain membranes with an $\alpha 1$ -subunit-specific antiserum. The receptors from $\alpha 1$ (H101R) mice displayed a ligand-binding profile consistent with that of physiologically diazepam-insensitive GABA_A receptors¹⁰, that is, a virtual lack of affinity for diazepam, clonazepam and zolpidem (Fig. 2c). The diazepam-insensitive sites in $\alpha 1$ (H101R) mutant brain were visualized autoradiographically by [³H]Ro15-4513 binding in the presence of 100 μ M diazepam in all regions known to express the $\alpha 1$ -subunit, including olfactory bulb, cerebral cortex, thalamus, pallidum, midbrain and cerebellum (Fig. 2d). In line with the known abundance of GABA_A $\alpha 1$ receptors, the number of diazepam-insensitive binding sites ($B_{\max}$: [³H]Ro15-4513 binding) increased from 7% in wild-type mice to 66% in $\alpha 1$ (H101R) mice without any change in the dissociation constant (K_d). The gating properties of the point-mutated receptor were assessed in whole-cell patch-clamp recordings from acutely dissociated cerebellar Purkinje cells, in which $\alpha 1$ receptors predominate. Although the response to GABA (3 μ M) was indistinguishable in cells from $\alpha 1$ (H101R) and wild-type mice, potentiation by diazepam (1 μ M) was greatly reduced in cells from $\alpha 1$ (H101R) mice (Fig. 2e). The remaining potentiation is attributed to diazepam-sensitive receptors other than $\alpha 1$ in these cells. The response to the inverse agonist Ro15-4513 (1 μ M) in wild-type cells was switched to an agonistic response in the mutant cells (Fig. 2e), which is in line with its agonistic effect at recombinant $\alpha 1$ (H101R) receptors⁷.

Diazepam and other classical benzodiazepines are widely used clinically as anxiolytics, hypnotics, anticonvulsants and myorelaxants, with side effects including anterograde amnesia, impairment of motor coordination and potentiation of ethanol effects. To identify the role of GABA_A $\alpha 1$ receptors in the actions of diazepam, various drug-induced behavioural responses were compared in $\alpha 1$ (H101R) mice and wild-type mice. Diazepam depressed motor activity, a measure of sedation, in wild-type but not in $\alpha 1$ (H101R) mice ($F(3, 120) = 4.90$, $P < 0.01$) (Fig. 3a). $\alpha 1$ (H101R) mice were also resistant to the sedative effect of zolpidem¹¹ (60 mg kg⁻¹ orally (p.o.)), a ligand with preferential affinity for GABA_A $\alpha 1$ receptors. The sedative or hypnotic effects of drugs other than ligands of the benzodiazepine site were unaltered in $\alpha 1$ (H101R) mice. The neurosteroid 3 α -hydroxy-5 β -pregnan-20-one (100 mg kg⁻¹ p.o.)¹² decreased motor activity to the same extent in $\alpha 1$ (H101R) and wild-type mice. Similarly, sodium pentobarbital, at the minimally effective dose of 100 mg kg⁻¹ intraperitoneally (i.p.), induced loss of righting reflex with the same latency in both genotypes. These results indicate that the lack of drug-induced sedation in $\alpha 1$ (H101R) mice is restricted to benzodiazepine-site ligands.

The memory impairing effect of diazepam was analysed in the step-through passive avoidance paradigm following administration of the drug 30 min before the training session¹³. The latency for re-entering the dark compartment 24 h later was shortened in wild-type ($U = 16.5$, $P < 0.01$ compared with vehicle) but not in $\alpha 1$ (H101R) mice ($U = 48.0$, n.s.) (Fig. 3b). In a modified lick suppression paradigm¹⁴, the latency to drink the first 100 licks was measured before a training session of punished drinking under vehicle or diazepam treatment (10 mg kg⁻¹ i.p.) and reassessed three days later (recall session) under drug-free conditions. Wild-type mice trained under vehicle displayed an enhanced recall latency (median (range): 928s (64–1,200) compared with a training latency of 12s (10–99), $P < 0.05$, $n = 7$). In contrast, wild-type mice trained under diazepam showed a short recall latency (13s (12–1,200) compared with a training latency of 29s (12–200), $n = 6$), suggesting a drug-induced amnesia. In both vehicle- and diazepam-treated $\alpha 1$ (H101R) mice, the recall latency (vehicle 1,200s (12–1,200), $n = 7$; diazepam 1,200s (45–1,200), $n = 6$) was enhanced compared with the training latency (vehicle 34s (14–62); diazepam 11s (10–28), $P < 0.05$). All four groups of mice afflicted on

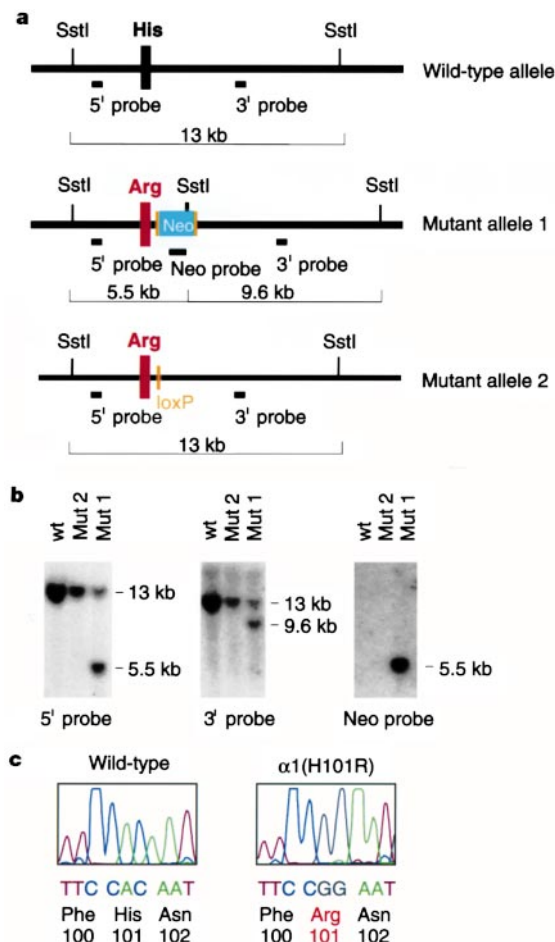


Figure 1 Targeting the GABA_A receptor $\alpha 1$ subunit (GABRA1) gene. **a**, The 5' and 3' flanking probes, which contain the genomic sequences directly adjacent to the targeting vector, are depicted as solid bars. His and Arg denote the codons for His 101 and Arg 101 in exon 4 of the wild-type and mutant alleles, respectively. **b**, Southern blot analysis of wild-type (wt) and mutant (Mut) alleles in mice. The genomic DNA was cut with *Sst*I. **c**, Verification of the $\alpha 1$ (H101R) point mutation by DNA sequencing. Exon 4 sequences were amplified from mouse genomic DNA by polymerase chain reaction and sequenced using an ABI Prism 310 Genetic Analyzer.

themselves a similar amount of punished licks during the training session, making a potential analgesic effect of diazepam unlikely. Thus, two different paradigms point to the absence of the amnesic effect of diazepam in $\alpha 1$ (H101R) mice. In contrast, the muscarinic antagonist, scopolamine (1.5 mg kg^{-1} i.p.), had an amnesic effect in both $\alpha 1$ (H101R) mice (latency $42.7 \pm 22.7 \text{ s}$ compared with vehicle: $165.8 \pm 14.2 \text{ s}$, $n = 5-8$, $P < 0.01$) and wild-type mice (latency $50.2 \pm 26.8 \text{ s}$ compared with vehicle: $180.0 \pm 0 \text{ s}$, $n = 4-8$, $P < 0.05$) in the passive avoidance protocol. These results underline

the specificity of the GABA_A $\alpha 1$ receptor in mediating diazepam-induced amnesia.

The anticonvulsant activity of diazepam was assessed against pentylenetetrazole-induced tonic convulsions¹⁵. It was reduced in $\alpha 1$ (H101R) mice ($\chi^2(4) = 12.50$, $P < 0.05$) compared with wild-type mice ($\chi^2(4) = 27.76$, $P < 0.001$). Indeed, 30 mg kg^{-1} diazepam protected all wild-type mice whereas 45.5% of $\alpha 1$ (H101R) mice still developed tonic seizures (Fig. 3c). The anticonvulsant effect of diazepam which remained in $\alpha 1$ (H101R) mice was due to GABA_A

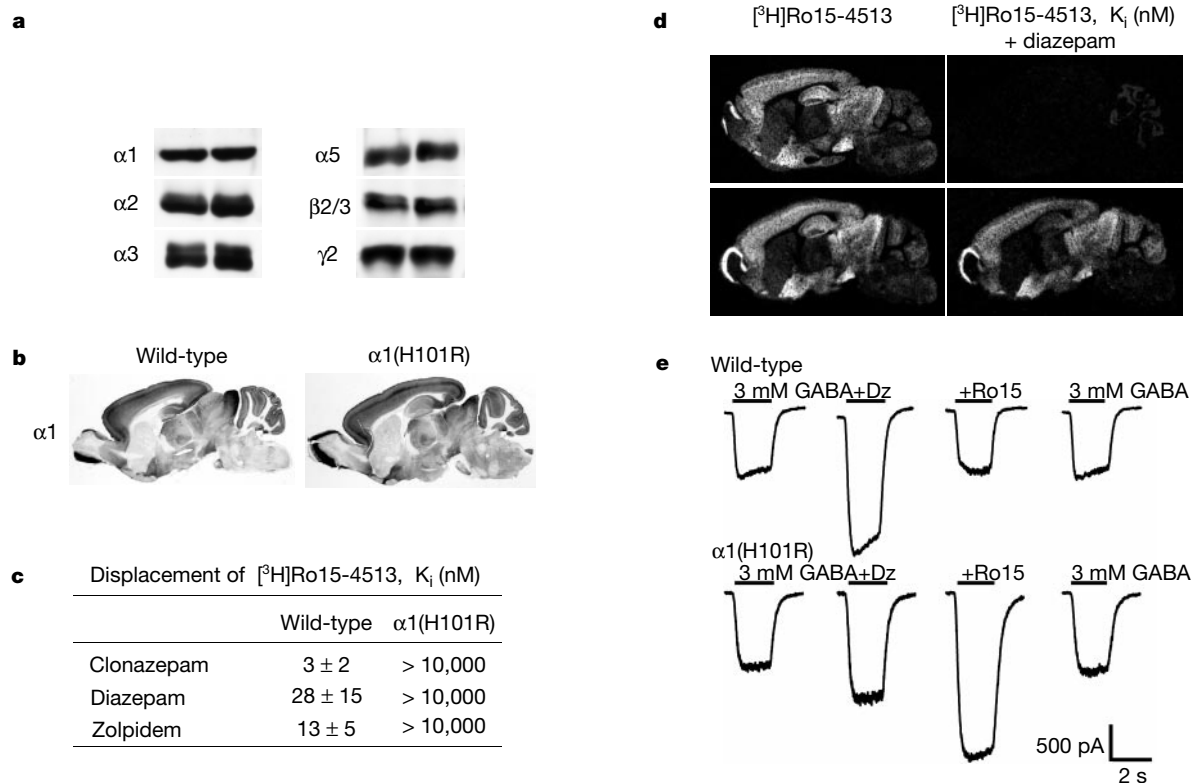


Figure 2 Biochemical, morphological and electrophysiological characteristics of GABA_A receptors in $\alpha 1$ (H101R) mice. **a**, Western blots of GABA_A receptor subunit proteins in wild-type and $\alpha 1$ (H101R) mice with antisera recognizing GABA_A receptor subunits $\alpha 1$, $\alpha 2$, $\alpha 3$, $\alpha 5$, $\beta 2/3$ and $\gamma 2$. **b**, Immunohistochemical regional distribution of the $\alpha 1$ subunit in parasagittal brain sections from wild-type (left) and $\alpha 1$ (H101R) mice (right). **c**, Displacement potencies of several benzodiazepine-site ligands at wild-type and $\alpha 1$ (H101R) receptors, immunoprecipitated with an $\alpha 1$ -subunit-specific antiserum. **d**, Autoradiographic distribution of total [^3H]Ro15-4513-binding sites (left) and diazepam-insensitive [^3H]Ro15-4513-binding sites (right) in wild-type mice (top) and $\alpha 1$ (H101R)

mice (bottom). Parasagittal brain sections were incubated with 20 nM [^3H]Ro15-4513 in the absence (left) or presence of $100 \mu\text{M}$ diazepam (right). Low levels of diazepam-insensitive binding in wild-type mice represent the receptors containing the $\alpha 4$ or $\alpha 6$ subunit. **e**, Patch-clamp analysis of GABA responses in dissociated cerebellar Purkinje cells. The potentiation by diazepam (Dz) was $134 \pm 19\%$ (mean $\pm$ s.e.m.; $n = 12$) in cells of wild-type mice and was reduced to $71 \pm 14\%$ ($n = 10$) in cells of $\alpha 1$ (H101R) mice ($P < 0.05$, t -test). The inverse agonist Ro15-4513 (Ro15) displayed an agonistic response in the mutant cells.

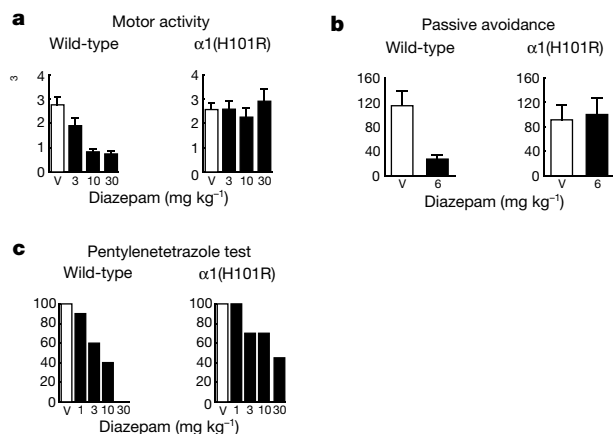


Figure 3 Behavioural assessment of drug-induced sedation, amnesia and anticonvulsant activity in wild-type and $\alpha 1$ (H101R) mice. **a**, Diazepam induced a dose-dependent decrease in motor activity in wild-type but not in $\alpha 1$ (H101R) mice ($n = 16$ per group). **b**, In a passive avoidance paradigm, diazepam shortened the latency to re-enter the dark compartment 24 h after training in wild-type mice but not in $\alpha 1$ (H101R) mice ($n = 10$ per group). **c**, Partial anticonvulsant activity of diazepam against pentylenetetrazole (120 mg kg^{-1} i.p.) in $\alpha 1$ (H101R) compared with wild-type mice ($n = 10-11$ per group). All results are given as means $\pm$ s.e.m. Asterisk, $P < 0.05$; double asterisks, $P < 0.01$; triple asterisks, $P < 0.001$. Diazepam was administered p.o.; V, vehicle.

receptors other than $\alpha 1$, as it was antagonized by the benzodiazepine antagonist, flumazenil¹⁶. Furthermore, sodium phenobarbital was fully effective as an anticonvulsant in $\alpha 1$ (H101R) mice with a dose–response relationship similar to that of wild-type mice. These results show that the anticonvulsant activity of benzodiazepine site ligands is partly mediated by GABA_A $\alpha 1$ receptors.

In striking contrast, the myorelaxant, motor-impairing, ethanol-potentiating and anxiolytic-like properties of diazepam were not impaired in the $\alpha 1$ (H101R) mice, indicating that they may be exclusively mediated by GABA_A receptors of the $\alpha 2$, $\alpha 3$ and/or $\alpha 5$ type. Diazepam reduced muscle tone to the same extent in wild-type ($\chi^2(4) = 32.24$, $P < 0.001$) and $\alpha 1$ (H101R) mice ($\chi^2(4) = 33.17$, $P < 0.001$) as indicated by a dose-dependent increase in the number of mice with impaired grasping reflex in the horizontal wire test¹⁵ (Fig. 4a). This myorelaxant effect was antagonized by flumazenil. In the rotarod test¹⁵, both $\alpha 1$ (H101R) and wild-type mice displayed a dose-dependent motor impairment (Treatment, $F(2, 54) = 14.82$, $P < 0.001$, analysis of variance (ANOVA) with repeated measures on the same subjects) (Fig. 4b). Furthermore, diazepam potentiated

Table 1 Roles of GABA_A receptor subtypes in benzodiazepine action

	$\alpha 1$	$\alpha 2, \alpha 3, \alpha 5$
Sedation	+	–
Amnesia	+	–
Seizure protection	+	+
Anxiolysis	–	+
Myorelaxation	–	+
Motor impairment	–	+
Ethanol potentiation	–	+

in a dose-dependent manner the sedative effect of ethanol by increasing the duration of the loss of the righting reflex¹⁷ in both wild-type and $\alpha 1$ (H101R) mice (Treatment, $F(3, 76) = 12.10$, $P < 0.001$) (Fig. 4c). Finally, the anxiolytic-like activity of diazepam appeared unaltered, as assessed in two classical paradigms. In the light/dark choice test¹⁸, diazepam dose-dependently increased the time spent in the illuminated area in wild-type and $\alpha 1$ (H101R) mice (Treatment, $F(2, 54) = 10.55$, $P < 0.001$) (Fig. 4d) without affecting the number of transitions. Likewise, in the elevated X-maze¹⁹, diazepam significantly increased the number of open arm entries for both wild-type and $\alpha 1$ (H101R) mice ($F(3, 68) = 3.93$, $P < 0.05$) (Fig. 4e). The increase in open arm entries seen in $\alpha 1$ (H101R) mice up to 2 mg kg^{-1} was maintained only up to 1 mg kg^{-1} in wild-type mice, presumably because of the susceptibility to sedation (Fig. 3a) of wild-type but not $\alpha 1$ (H101R) mice.

Our finding that benzodiazepine actions are mediated by specific GABA_A receptor subtypes points to new strategies for drug design for several therapeutic indications. For example, agonists acting on $\alpha 2$, $\alpha 3$, and/or $\alpha 5$ but not on $\alpha 1$ receptors are expected to be non-sedative and non-amnesic anxiolytics (Table 1). Furthermore, the $\alpha 1$ (H101R) mice provide a suitable model to investigate whether the development of tolerance and dependence liability of benzodiazepines is also subtype specific. In addition, our results are likely to be of physiological relevance, as benzodiazepine agonists act by enhancing physiological GABA-mediated transmission. It is therefore conceivable that the physiological regulation of vigilance and memory function is mediated by circuits involving GABA-mediated transmission at $\alpha 1$ receptors, whereas neurons with benzodiazepine-sensitive GABA_A receptors other than $\alpha 1$ are involved in regulating anxiety. Thus, GABA_A receptor subtypes promise to define circuits mediating distinct behaviours. □

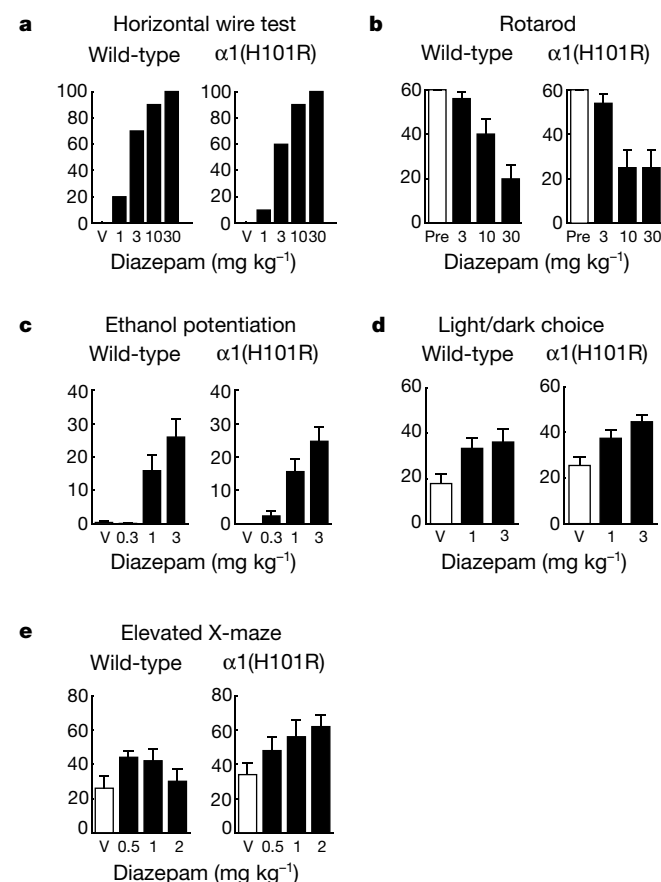


Figure 4 Behavioural assessment of myorelaxant, motor-impairing, ethanol-potentiating and anxiolytic-like actions of diazepam in wild-type and $\alpha 1$ (H101R) mice. **a**, Dose-dependent impairment of the grasping reflex in wild-type and $\alpha 1$ (H101R) mice ($n = 10$ – 11 per group). **b**, Dose-dependent decrease of the latency to fall off the rotating rod in wild-type and $\alpha 1$ (H101R) mice ($n = 10$ per group). **c**, Dose-dependent potentiation of the sedative effect of a 15% ethanol solution in wild-type and $\alpha 1$ (H101R) mice ($n = 5$ – 15 per group). **d**, Dose-dependent increase of the percentage of time spent in the lit area in wild-type and $\alpha 1$ (H101R) mice ($n = 10$ per group) tested in a light/dark choice paradigm. **e**, Dose-dependent increase of the number of open arm entries (in percentage of total number of arm entries) in both wild-type and $\alpha 1$ (H101R) mice ($n = 8$ – 11 per group) exposed to an elevated X-maze. Results are given as means $\pm$ s.e.m. Asterisks, $P < 0.05$; double asterisks, $P < 0.01$; triple asterisks, $P < 0.001$. Diazepam was administered p.o.; V, vehicle.

Methods

Strain background of mice

$\alpha 1$ (H101R) chimera mice were crossed first to Ella-cre mice (third backcross to 129/SvEv), to remove the neo cassette from the germline⁹ through *cre/loxP*-mediated excision, twice more to 129/SvEv mice and then for five generations to both 129/SvEv and C57BL/6J. Heterozygote crosses were set up for each strain to yield homozygous mutant, heterozygous and wild-type mice. Typically, 20–25 breeding pairs of both homozygous mutant and wild-type mice produced the experimental animals which were used at roughly 8–12 weeks of age. Each mouse was injected with diazepam only once.

Behavioural tests

All tests were performed in two backgrounds: motor activity, pentylentetrazole test, horizontal wire test, rotarod and ethanol potentiation in 129/SvEv (Fig. 3a, c, d) and 129/SvJ (Fig. 4b); passive avoidance and elevated X-maze in 129/SvJ (Fig. 3b) and C57BL/6J (Fig. 4e); light/dark choice test in 129/SvJ (Fig. 4d) and in 129/SvJ $\times$ C57BL/6J F1 hybrids. The 129/SvEv background was used for binding studies, autoradiography, immunoprecipitation and immunohistochemistry, the 129/SvJ background for electrophysiology.

Received 2 June; accepted 31 August 1999.

1. Fritschy, J. M. & Mohler, H. GABA_A-receptor heterogeneity in the adult rat brain: differential regional and cellular distribution of seven major subunits. *J. Comp. Neurol.* **359**, 154–194 (1995).
2. Barnard, E. A. *et al.* International Union of Pharmacology. XV. Subtypes of γ -aminobutyric acid_A receptors: classification on the basis of subunit structure and receptor function. *Pharmacol. Rev.* **50**, 291–313 (1998).

- Mohler, H. & Okada, T. Benzodiazepine receptor: demonstration in the central nervous system. *Science* **198**, 849–851 (1977).
- Braestrup, C., Albrechtsen, R. & Squires, R. F. High densities of benzodiazepine receptors in human cortical areas. *Nature* **269**, 702–704 (1977).
- Wieland, H. A., Lüddens, H. & Seeburg, P. H. A single histidine in GABA_A receptors is essential for benzodiazepine agonist binding. *J. Biol. Chem.* **267**, 1426–1429 (1992).
- Kleingor, C., Wieland, H. A., Korpi, E. R., Seeburg, P. H. & Kettenmann, H. Current potentiation by diazepam but not GABA sensitivity is determined by a single histidine residue. *Neuroreport* **4**, 187–190 (1993).
- Benson, J. A., Löw, K., Keist, R., Mohler, H. & Rudolph, U. Pharmacology of recombinant γ -aminobutyric acid_A receptors rendered diazepam-insensitive by point-mutated α -subunits. *FEBS Lett.* **431**, 400–404 (1998).
- Fritschy, J.-M. *et al.* Five subtypes of type A γ -aminobutyric acid receptors identified in neurons by double and triple immunofluorescence staining with subunit-specific antibodies. *Proc. Natl Acad. Sci. USA* **89**, 6726–6730 (1992).
- Lakso, M. *et al.* Efficient *in vivo* manipulation of mouse genomic sequences at the zygote stage. *Proc. Natl Acad. Sci. USA* **93**, 5860–5865 (1996).
- Knoflach, F. *et al.* Pharmacological modulation of the diazepam-insensitive recombinant γ -aminobutyric acid_A receptors $\alpha 4\beta 2\gamma 2$ and $\alpha 6\beta 2\gamma 2$. *Mol. Pharmacol.* **50**, 1253–1261 (1996).
- Bernard, F. *et al.* in *Zolpidem: an Update of its Pharmacological Properties and the Therapeutic Place in the Management of Insomnia* (eds Freeman, H., Puech, A. J. & Roth, T.) 21–31 (Elsevier, Paris, 1996).
- Melchior, C. L. & Allen, P. M. Interaction of pregabalin and pregabalin sulfate with ethanol and pentobarbital. *Pharmacol. Biochem. Behav.* **42**, 605–611 (1992).
- Beuzen, A. & Belzung, C. Link between emotional memory and anxiety states: a study by principal component analysis. *Physiol. Behav.* **58**, 111–118 (1995).
- Vogel, J. R., Beer, B. & Clody, D. E. A simple and reliable conflict procedure for testing anti-anxiety agents. *Psychopharmacology* **21**, 1–7 (1971).
- Bonetti, E. P. *et al.* Benzodiazepine antagonist Ro 15-1788: neurological and behavioral effects. *Psychopharmacology* **78**, 8–18 (1982).
- Hunkeler, W. *et al.* Selective antagonists of benzodiazepines. *Nature* **290**, 514–516 (1981).
- Bonetti, E. P. *et al.* Ro 15-4513: Partial inverse agonism at the BZR and interaction with ethanol. *Pharmacol. Biochem. Behav.* **31**, 733–749 (1988).
- Misslin, R., Belzung, C. & Vogel, E. Behavioural validation of a light/dark choice procedure for testing anti-anxiety agents. *Behav. Proc.* **18**, 119–132 (1989).
- Lister, R. G. The use of a plus-maze to measure anxiety in the mouse. *Psychopharmacology* **92**, 180–185 (1987).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

Acknowledgements

We thank Y. Lang for blastocyst injections, D. Blaser, S. Ganz, H. Pochetti and G. Schmid for animal care, R. Keist and C. Michel for technical assistance and H. Westphal for providing Ella-cre mice. This work was supported by a grant from the Swiss National Science Foundation.

Correspondence and requests for materials should be addressed to U.R. (e-mail: rudolph@pharma.unizh.ch).

P/Q-type calcium channels mediate the activity-dependent feedback of syntaxin-1A

Kathy G. Sutton^{*†}, John E. McRory^{*†}, Heather Guthrie^{*}, Timothy H. Murphy[‡] & Terrance P. Snutch^{*}

^{*} Biotechnology Laboratory and [‡] Kinsmen Laboratory of Neurological Research, Dept Psychiatry University of British Columbia, Vancouver, V6T 1Z3, Canada
[†] These authors contributed equally to this work.

Spatial and temporal changes in intracellular calcium concentrations are critical for controlling gene expression in neurons^{1–5}. In many neurons, activity-dependent calcium influx through L-type channels stimulates transcription that depends on the transcription factor CREB by activating a calmodulin-dependent pathway^{6–11}. Here we show that selective influx of calcium through P/Q-type channels^{12–14} is responsible for activating expression of syntaxin-1A, a presynaptic protein that mediates vesicle docking, fusion and neurotransmitter release. The initial P/Q-type calcium signal is activated by release of calcium from intracellular stores

and acts through phosphorylation that is dependent on the calmodulin-dependent kinase CaM K II/IV, protein kinase A and mitogen-activated protein kinase kinase. Initiation of syntaxin-1A expression is rapid and short-lived, with syntaxin-1A ultimately interacting with the P/Q-type calcium channel to decrease channel availability. Our results define an activity-dependent feedback pathway that may regulate synaptic efficacy and function in the nervous system.

Syntaxin-1A interacts with the synaptic core (SNARE) complex and is tightly regulated by a number of proteins including munc-13, munc-18 and tomosyn^{15–18}. Syntaxin-1A also binds to the domain II–III linker of N- and P/Q-type voltage-gated Ca²⁺ channels and shifts current inactivation properties to more negative potentials^{19–23}. Transient expression of human α_{1A} (P/Q-type) Ca²⁺ channels in HEK293 cells resulted in a relatively negative steady-state inactivation (SSI) and indicated that there may be a corresponding interaction with syntaxin (Fig. 1a; voltage for 50% inactivation ($V_{50\text{inact}}$) = -69.2 ± 1.6 mV, $n = 17$). Incubation with botulinum toxin C1 (200–400 ng ml⁻¹ for 24 h) to cleave syntaxin²⁴ made the SSI ~ 10 mV more depolarized (Fig. 1a; $V_{50\text{inact}}$ = -59.7 ± 1.7 mV, $P < 0.001$, $n = 11$). Treatment with botulinum C1 toxin did not affect the voltage-dependence of current activation (Fig. 1a inset). To confirm the presence of syntaxin, cells were co-transfected with an antisense syntaxin construct directed against the first 338 bases of rat syntaxin (GenBank accession code M95734; 90% identical with human syntaxin-1A), which resulted in a depolarizing shift in the SSI curve (Fig. 1b; $V_{50\text{inact}}$ = -63.1 ± 1.3 mV, $P < 0.01$, $n = 18$). Thus, syntaxin-1A endogenous to HEK293 cells interacts with transiently expressed α_{1A} P/Q-type Ca²⁺ channels and shifts the channels into a more inactivated state.

Unexpectedly, western blot analysis using a monoclonal antibody against human syntaxin-1A²⁵ showed that syntaxin-1A was not endogenously expressed in HEK293 cells (Fig. 2a); however, a 34–36 kilodalton band that co-migrated with syntaxin-1A protein was detected in HEK293 cells transfected with the α_{1A} P/Q-type Ca

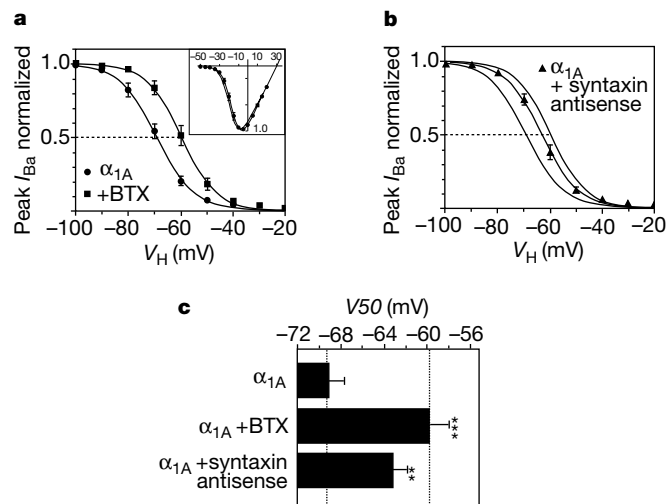


Figure 1 Interaction of syntaxin with α_{1A} P/Q-type channels induces a negative shift in steady-state inactivation (SSI). **a**, Cells were held at the holding voltage (V_h) for 15 s before activation; command voltage (V_c) = 0 mV. Data were normalized to the maximum current ($V_h = -120$ mV) and fitted using $y = 1/(1 + \exp((V_{50\text{inact}} - x)/k))$. Botulinum toxin (BTX) shifts $V_{50\text{inact}}$ to a more depolarized value without any change in the slope ($k = -5.6 \pm 0.2$, $P = 0.42$). **b**, Antisense syntaxin-1A induces a depolarizing shift in the SSI of α_{1A} . Dashed lines represent fits in **(a)**. **c**, Summary of mean ($\pm$ s.e.m.) $V_{50\text{inact}}$. ** = $P < 0.01$ and *** = $P < 0.001$ compared with α_{1A} control. I_{Ba} , maximum peak current.

channel. Polymerase chain reaction of reverse transcribed RNA (RT-PCR) also showed that syntaxin-1A messenger RNA was undetectable in untransfected HEK293 cells. In contrast, 24–36 h after transfection with α_{1A} , syntaxin-1A mRNA appeared and paralleled the expression of the exogenous P/Q-type Ca^{2+} channel (Fig. 2b). Syntaxin-1A mRNA was inhibited in transfected cells incubated in actinomycin D ($2 \mu\text{g ml}^{-1}$; data not shown), indicating that the P/Q-type response may involve transcriptional induction, although we cannot rule out effects on mRNA stability²⁶. Co-transfection with only Ca^{2+} channel $\alpha_2\delta$ and β_{1b} subunits, CD8 vector and Bluescript plasmid did not induce syntaxin-1A expression (data not shown). Similarly, in α_{1A} -transfected cells RT-PCR did not show the expression of other SNARE proteins including syntaxin 1B, SNAP-25, synaptophysin, VAMP and synaptotagmin (data not shown).

To test whether functional P/Q-type Ca^{2+} channels were required to induce syntaxin-1A, we examined the effects of the selective P/Q-type channel antagonist ω -Agatoxin IVA (ω -Aga IVA). Figure 2c shows that syntaxin-1A mRNA was not induced in the presence of ω -Aga IVA (20 nM). Similar results were obtained when ω -Aga IVA

was applied either at the time of α_{1A} transfection or 24 h after transfection (data not shown). Syntaxin-1A mRNA was also undetectable when ω -CTX MVIIC (200 nM) was applied just 4 h before cell harvesting, which implies the rapid turnover of syntaxin-1A mRNA (Fig. 2c). Furthermore, the presence of functional Ca^{2+} channels *per se* was not sufficient, as the transfection of other cloned Ca^{2+} channels, including α_{1B} (N-type), α_{1C} (L-type), α_{1E} (novel type) and the α_{1G} and α_{1I} T-type channels, did not result in detectable levels of syntaxin-1A mRNA (Fig. 2c). Stimulation of α_{1C} L-type Ca^{2+} channels with either 40 mM KCl or BAY-K8644 (10 μM) also failed to activate production of syntaxin-1A mRNA (data not shown). Thus, our data indicate that the selective functional expression of P/Q-type Ca^{2+} channels mediates the induction of syntaxin-1A.

The resting membrane potential of HEK293 cells in culture could conceivably allow a window current of Ca^{2+} influx through P/Q channels. Fura-2-based Ca^{2+} imaging showed that HEK293 cells transfected with α_{1A} Ca^{2+} channels (>48 h) exhibit elevated resting Ca^{2+} levels when compared with untransfected cells

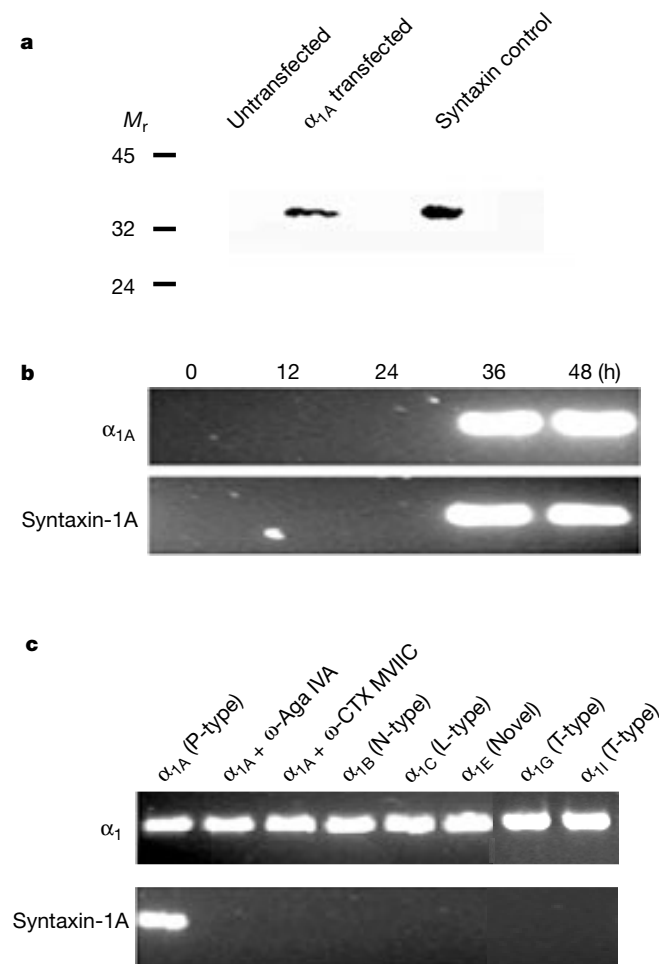


Figure 2 Functional α_{1A} P/Q channels stimulate the expression of syntaxin-1A. **a**, Western blot of control and transfected HEK293 cells shows syntaxin-1A protein after expression of α_{1A} channels. **b**, RT-PCR of HEK293 cells at increasing time intervals after transfection. Sequence analysis of the gel-eluted PCR product confirmed the identification of the PCR product as the human isoform syntaxin-1A mRNA (accession code L37792). **c**, The presence of either ω -Aga IVA or ω -CTX MVIIC during transfection and incubation of the cells prevents syntaxin-1A expression. Expression of α_{1B} , α_{1C} , α_{1E} , α_{1G} or α_{1I} Ca^{2+} -channel subtypes does not induce syntaxin-1A mRNA expression. M_r , relative molecular mass ($\times 10^3$).

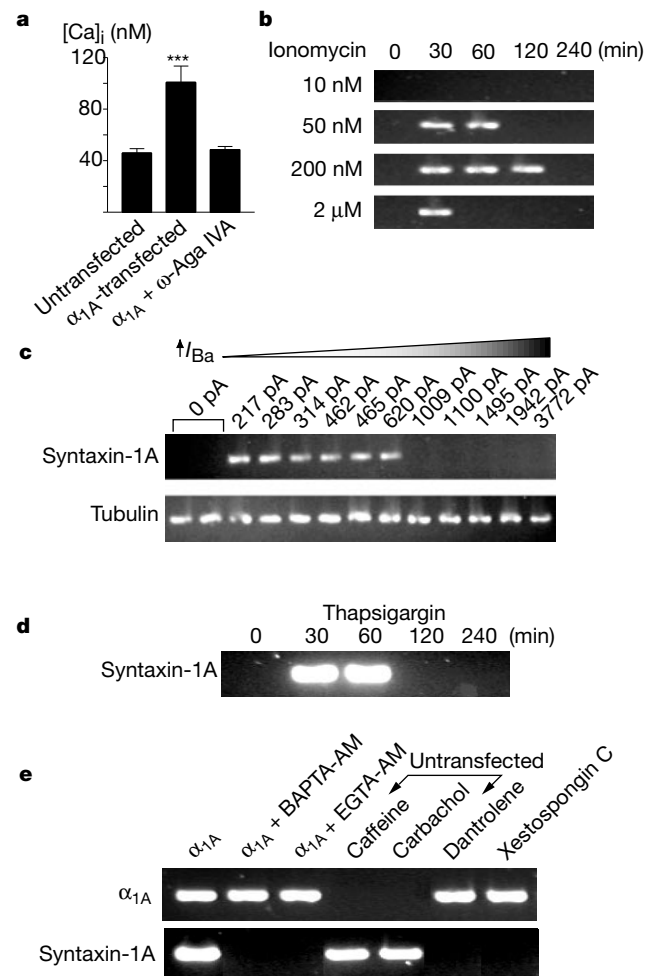


Figure 3 Syntaxin-1A expression is stimulated by distinct levels of intracellular Ca^{2+} and is mediated through activation of Ca^{2+} stores. **a**, $[\text{Ca}]_i$ levels are elevated in α_{1A} -transfected HEK293 cells; *** $P < 0.0001$. **b**, Ionomycin-induced Ca^{2+} influx activates syntaxin-1A expression in untransfected cells. **c**, Single-cell PCR and whole-cell P/Q-type currents of individual α_{1A} transfected cells. **d**, Thapsigargin transiently induces syntaxin-1A expression. **e**, Loading with BAPTA-AM or EGTA-AM (30 min at 2 μM) 4 h before harvesting blocks syntaxin-1A expression. Syntaxin-1A expression is stimulated in untransfected cells by release of Ca^{2+} from Ca^{2+} stores and blocked in α_{1A} -transfected cells when Ca^{2+} stores are inhibited.

(100.8 ± 12.6 nM and 46.1 ± 3.3 nM, respectively; $n = 67, 54$; Fig. 3a). Moreover, intracellular Ca^{2+} concentration returned to control levels when α_{1A} -transfected cells were incubated with ω -Aga IVA (48.7 ± 2.4 nM, $n = 34$), confirming that cultured HEK293 cells are subject to a basal Ca^{2+} current through exogenously expressed P/Q-type Ca^{2+} channels.

The requirement for extracellular Ca^{2+} was also investigated in untransfected cells using a 10-min application of ionomycin to produce a transient increase in intracellular Ca^{2+} . Applications of different concentrations of ionomycin (10 nM–2 μ M) determined that syntaxin-1A expression exhibited a discrete upper and lower limit of Ca^{2+} -dependent activation with maximal activation occurring at relatively low levels of ionomycin (50–200 nM; Fig. 3b). Conversely, induction of syntaxin-1A mRNA in transfected cells was blocked when external Ca^{2+} was buffered using EGTA (2 mM; data not shown). The basal Ca^{2+} influx in α_{1A} -transfected cells raised the issue of whether the P/Q-type specificity of the syntaxin response was due to differences in expression levels for the different types of Ca^{2+} channel examined. Equivalent expression levels (measured as peak current amplitudes) for α_{1A} , α_{1B} and α_{1C} were investigated on a cell-by-cell basis and correlated with the presence or absence of syntaxin-1A mRNA using single cell RT-PCR. Figure 3c shows that only cells transfected with α_{1A} that exhibited currents within the range of ~200–600 pA induced syntaxin-1A ($n = 6$). In contrast, syntaxin-1A mRNA was undetectable in all the α_{1B} (N-type) and α_{1C} (L-type) transfected cells (peak current amplitudes $\approx$ 200–600 pA, $n = 4$ and 5, respectively; data not shown).

To map the intracellular regions that are effective in transducing the α_{1A} -specific Ca^{2+} influx, we tested a number of agents that alter intracellular Ca^{2+} levels ($[\text{Ca}]_i$). Release of Ca^{2+} from intracellular stores with a 10-min application of thapsigargin (10 μ M) stimulated syntaxin-1A expression in untransfected cells (Fig. 3d). In α_{1A} -transfected cells both BAPTA-AM and EGTA-AM blocked the Ca^{2+} -dependent induction of syntaxin-1A (Fig. 3e). The slower Ca^{2+} -

binding kinetics of EGTA (compared with BAPTA) limits its effective buffering capacity to regions more distal to the voltage-dependent Ca^{2+} influx⁷. Application of either caffeine (10 mM) or carbachol (20 μ M) to activate Ca^{2+} release from ryanodine- and D-myo-inositol 1,4,5-trisphosphate (IP_3)-sensitive stores rapidly stimulated (<30 min) syntaxin-1A mRNA levels in untransfected cells (Fig. 3e). Conversely, block of either ryanodine-sensitive Ca^{2+} release with dantrolene (10 μ M) or IP_3 -induced Ca^{2+} release with Xestospongion C (1 μ M)²⁷ inhibited α_{1A} -specific syntaxin-1A expression in transfected cells. Thus, a discrete level of basal α_{1A} -specific Ca influx induces syntaxin-1A expression through a secondary Ca^{2+} release from intracellular stores.

Ca^{2+} -dependent messengers and transcription factors such as CREB are important in initiating gene transcription^{1,2}. Several different Ca^{2+} -dependent kinases and phosphatases, including calmodulin/ Ca^{2+} -dependent kinase (CaM) kinases I, II and IV, protein kinase A (PKA) and the Ras mitogen-activated kinase (MAPK) signalling pathway are responsible for regulating CREB activation^{2,28,29}. Incubation of α_{1A} -transfected cells with calmidazolium (10 μ M) to block calmodulin activity or KN-62 (900 nM), an inhibitor of CaM kinases II and IV, prevented syntaxin-1A expression (Fig. 4a). Incubation with H-9 (20 μ M), H-8 (1 μ M) or Rp-cAMP (10 μ M) to inhibit PKA resulted in the block of syntaxin-1A expression in α_{1A} -transfected cells (Fig. 4a). Conversely, a 10-min application of the PKA activators forskolin (20 μ M) or Sp-cAMPS (10 μ M) induced syntaxin-1A mRNA in the absence of α_{1A} P/Q-type channels. Activation of protein kinase C with 12-O-tetradecanoylphorbol-13-acetate (TPA) did not induce expression in the absence of α_{1A} channels. Lastly, block of tyrosine kinase and MAPK kinase (MAPKK) activities by incubation with genistein (50 μ M), and PD 098059 (50 μ M), respectively, inhibited syntaxin-1A expression in α_{1A} -transfected cells, indicating that there may be involvement of the rsk kinase pathway.

Western blot analysis of control and α_{1A} -transfected cells showed

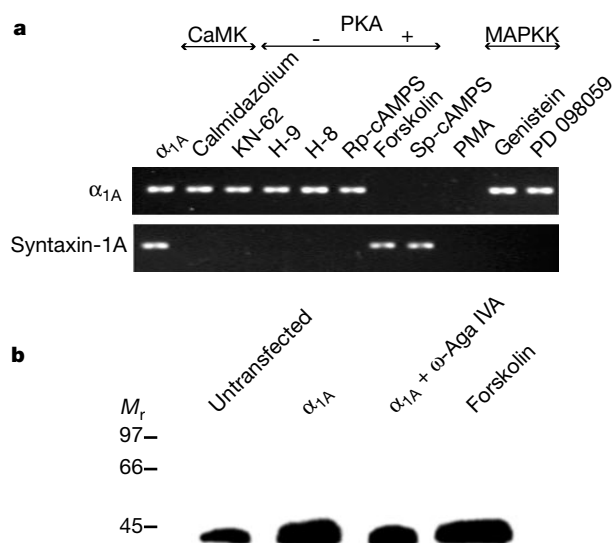


Figure 4 Involvement of second messengers in syntaxin-1A expression. **a**, Inhibition of calmodulin, CaM kinase II/IV or PKA blocks syntaxin-1A expression in α_{1A} -transfected cells. Stimulation of PKA with forskolin (20 μ M) or Sp-cAMPS (10 μ M) in untransfected cells activates expression (1 h after treatment). Activation of PKC with PMA (100 nM, 1 h) is ineffective, whereas incubation with genistein (50 μ M) or PD 098059 blocks expression. **b**, Western blot for phosphorylated CREB. Lane 1, control HEK293 cell lysate; lane 2, α_{1A} -transfected HEK293 cells; lane 3, α_{1A} transfected HEK293 cells + ω -Aga-IVA (20 nM, 10 h); lane 4, control HEK293 cells + 20 μ M forskolin (30 min).

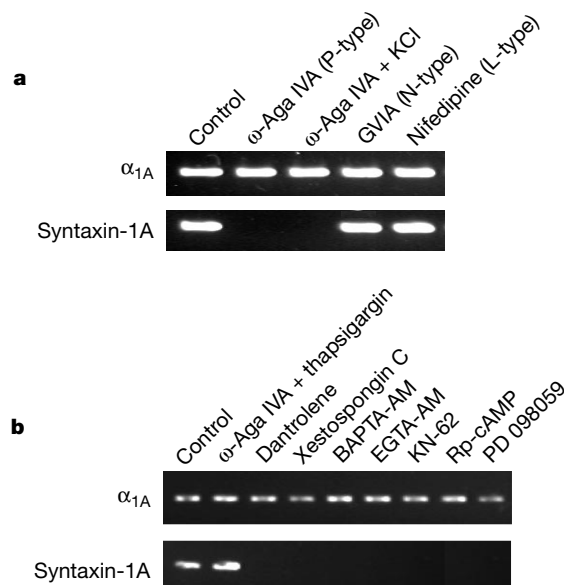


Figure 5 P/Q-type channels trigger syntaxin-1A expression in cerebellar granule neurons. **a**, RT-PCR from cultured cerebellar granule-cells. Block of P/Q-type channels (ω -Aga IVA, 100 nM) inhibits syntaxin-1A expression even in the presence of 40 mM KCl (30 min); inhibition of N-type channels with ω -CTx GVIA (1 μ M) or L-type channels with nifedipine (10 μ M) has no effect. **b**, Thapsigargin applied to granule cells (10 μ M, 10 min) incubated in ω -Aga IVA (100 nM) reactivates syntaxin-1A expression. Dantrolene (10 μ M), Xestospongion C (1 μ M), BAPTA-AM (2 μ M), EGTA-AM (2 μ M), KN-62 (900 nM), Rp-cAMPS (10 μ M) and PD 098059 (50 μ M) block expression.

that the level of phosphorylated CREB is upregulated in HEK293 cells expressing functional α_{1A} Ca^{2+} channels (Fig. 4b). Moreover, the level of CREB phosphorylation is dependent upon Ca^{2+} influx through functional α_{1A} channels, as it returns to basal levels in the presence of ω -Aga IVA. Therefore, syntaxin-1A expression may be mediated by CREB, although the involvement of additional cofactors, such as CREB-binding protein, or Ca^{2+} -dependent transcription factors, such as DREAM, cannot be ruled out^{2,10,11,30}.

To correlate the results in transfected HEK293 cells with P/Q-type channels in neurons, syntaxin-1A expression was investigated in cultured cerebellar granule cells. RT-PCR of granule neurons showed that syntaxin-1A is basally expressed (Fig. 5a). Blockade of P/Q-type channels by incubation with ω -Aga IVA (100 nM) reduced syntaxin mRNA to undetectable levels, even in the presence of 40 mM KCl to maximize activation of Ca^{2+} channels insensitive to ω -Aga IVA. Moreover, inhibition of N- or L-type current with ω -CTx GVIA or nifedipine was ineffective at blocking the Ca^{2+} -dependent expression (Fig. 5a). Application of thapsigargin to granule cells incubated in ω -Aga IVA recovered syntaxin-1A expression (Fig. 5b). Furthermore, expression of syntaxin-1A was inhibited by blocking either ryanodine- or IP_3 -induced Ca^{2+} release from stores, or by inhibiting CaM Kinase II/IV, PKA and MAPKK in a similar manner to that in HEK293 cells (Fig. 5b).

It is well established that Ca^{2+} influx through L-type Ca^{2+} channels stimulates CREB-activated gene expression^{6–11}. Our results show that discrete levels of Ca^{2+} influx through P/Q-type channels provides an alternative means of mediating Ca^{2+} -dependent gene expression. P/Q-type-selective Ca^{2+} influx induces the expression of syntaxin-1A and not of other SNARE proteins involved in vesicle fusion and neurotransmitter release. The regulation of syntaxin-1A appears to be under tight spatial and temporal Ca^{2+} -dependent control and is similar to that for the Ca^{2+} -dependent expression of other genes^{4,5}. This finding, coupled with a supporting role for Ca^{2+} release from intracellular stores, indicates that there may be a privileged close association between P/Q-type channels and Ca^{2+} stores. The induced syntaxin-1A protein appears to act as a negative feedback regulator of P/Q-type channels and may have a critical role in modulating synaptic activity. □

Methods

Electrophysiology and cell culture

HEK293 ts201 cells were transiently transfected with Ca^{2+} -channel subunits using standard Ca^{2+} -phosphate precipitation. Human α_{1A} (P/Q-type), rat α_{1B} (N-type), rat α_{1C} (L-type) and rat α_{1E} (novel type) Ca^{2+} -channel subunits (1–3 μ g of complementary DNA) were co-transfected with rat $\alpha_{1\delta}$ and β_{1B} ancillary subunits (molar ratio 1:1:1), 2 μ g of CD8 marker plasmid and 9–15 μ g Bluescript SK carrier DNA (total = 20 μ g cDNA per transfection). The α_{1G} and α_{1I} T-type Ca^{2+} -channel subunits were transfected without Ca^{2+} -channel ancillary subunits. Whole-cell patch clamp recordings were performed 48–96 h after transfection. The internal patch-pipette solution contained (in mM): 105 CsCl, 25 TEACl, 1 $CaCl_2$, 11 EGTA, 10 HEPES, pH 7.2. The external recording solution contained (in mM): 5 $BaCl_2$, 1 $MgCl_2$, 10 HEPES, 40 TEACl, 10 Glucose, 87.5 CsCl, pH 7.2. Currents were typically elicited from a holding potential of –120 mV to maximize the availability of syntaxin-1A-bound α_{1A} channels²⁰. Data were filtered at 1 kHz and in most cases subtraction of capacitance and leakage currents was carried out on-line using a P/4 protocol.

Cerebellar granule cells were prepared from 7-day rat pups by incubation of cerebellar pieces with trypsin solution (0.025% trypsin, 0.25% glucose, 0.3% BSA, 35 mM $MgSO_4$) at 37 °C for 20 min and treatment with DNase I (1 mg ml^{–1} in 0.1 mg ml^{–1} trypsin inhibitor and 35 mM $MgSO_4$); cells were collected by centrifugation. After trituration and pelleting, cells were incubated at 37 °C in MEM + 10% FCS for 24 h and then in MEM + 10% FCS + 5 fluoro 5'-deoxyuridine.

RNA preparation and RT-PCR

For each experiment, approximately 10⁶ HEK293 cells were harvested 48 h after transfection and resuspended in 1 ml of extraction buffer (4M guanidinium thiocyanate, 20 μ l β mercaptoethanol, 100 μ l of Na-acetate pH 4.0), 1 ml H₂O-saturated phenol and 200 μ l chloroform. Unless otherwise stated, all inhibitory agents were added 4 h before harvesting. After isolation of RNA, reverse transcriptase reactions were performed with 1 μ g total RNA and 20 pmol of a 3' oligonucleotide (antisense). The syntaxin-1A 3' oligonucleotide corresponds to basepairs (bp) 860–841 of the human syntaxin-1A mRNA sequence (accession code M95734). Other 3' antisense oligonucleotides utilized included

syntaxin 1B bp 749–769 (accession code 4507290); synaptophysin bp 919–902 (accession code X06389); synaptotagmin bp 1,783–1,765 (accession code M55047); VAMP bp 279–262 bp (accession code AF060538); SNAP-25 bp 453–436 (accession code L19760); α_{1A} Ca^{2+} channel bp 1,262–1,244 (accession code M64373); and α -tubulin bp 738–721 (accession code V01227). RNA and 3' oligonucleotides were heated to 90 °C for 10 min, placed on ice and a mixture of 1 $\times$ RT buffer, 10 mM dNTPs, 100 mM DTT and 5 units of Superscript RT (Gibco BRL) was added to a final volume of 20 μ l. Reactions were mixed and incubated at 42 °C for 1.5 h.

PCR was carried out with 5 μ l of the RT mix, 20 pmol of each 5' and 3' oligonucleotide, 10 mM dNTPs, 2.5 mM $MgCl_2$ and 2 units of TAQ polymerase to a final volume of 50 μ l. The human syntaxin-1A 5' oligonucleotide utilized corresponds to bp 241–258; syntaxin-1B to bp 342–362; synaptophysin to bp 556–574; synaptotagmin to bp 799–816; VAMP to bp 34–52; SNAP-25 to bp 89–106; α_{1A} P/Q-type to bp 652–670; and α -tubulin to bp 592–612. The PCR cycling parameters were 92 °C (1 min), 50 °C (1.25 min), 74 °C (1 min) for 10 cycles, followed by 92 °C (1 min), 57 °C (0.75 min), 74 °C (1 min) for 20 cycles. To confirm the identity of syntaxin-1A, PCR reactions were separated routinely through a 1.3% agarose gel, transferred to a Nytran membrane and probed with a radiolabelled syntaxin-1A-specific oligonucleotide.

For single-cell PCR experiments, the RT and PCR protocols were as described above except that cell contents were drawn into the recording pipette and added to 12 μ l of the RT mix without reverse transcriptase. One half of the RT mix was used for the syntaxin-1A PCR and the other half for either tubulin or Ca^{2+} -channel α_1 -subunit PCR.

Calcium imaging

Calcium imaging was performed using 17 μ M fura-2 AM loaded into HEK293 cells (30 min, room temperature in 0.067% pluronic). Resting $[Ca]_i$ was estimated by taking 380/360 nm excitation image pairs (isosbestic point ratio) in DMEM culture medium that was equilibrated with a 5% CO₂/95% air mixture and continuously perfused. Ratio values were converted to $[Ca^{2+}]_i$ values based on parameters obtained from an *in vitro* calibration.

Western blot

For each experiment, approximately 10⁶ HEK293 cells were harvested 48 h after transfection and lysed directly into loading buffer (4% SDS, 20% glycerol, 0.12 M Tris pH 6.8, 0.01% bromophenol blue, 0.1% DTT). Samples were boiled and proteins separated by electrophoresis through a 12% SDS-PAGE gel before transfer to a nylon filter (Nytran). Blots were probed with primary antibodies against the phosphorylated form of CREB (pCREB, Upstate Biotechnology) or human syntaxin (SP8, gift of W. Honer), and secondary antibody (anti-mouse Ig, Amersham) before bands were visualized using a commercial ECL detection kit (Amersham).

Received 4 May; accepted 26 August 1999.

- Ghosh, A. & Greenberg, M. E. Calcium signalling in neurons: molecular mechanisms and cellular consequences. *Science* **268**, 239–247 (1995).
- Bito, H., Deisseroth, K. & Tsien, R. W. Ca^{2+} -dependent regulation in neuronal gene expression. *Curr. Opin. Neurobiol.* **7**, 419–429 (1997).
- Hardingham, G. E., Chawla, S., Johnson, C. M. & Bading, H. Distinct functions of nuclear and cytoplasmic calcium in the control of gene expression. *Nature* **385**, 260–265 (1997).
- Dolmetsch, R. E., Lewis, R. S., Goodnow, C. C. & Healy, J. I. Differential activation of transcription factors induced by Ca^{2+} response amplitude and duration. *Nature* **386**, 855–858 (1997).
- Li, W., Llopis, J., Whitney, M., Zlokarnik, G. & Tsien, R. Y. Cell-permeant caged $InsP_3$ ester shows that Ca^{2+} spike frequency can optimize gene expression. *Nature* **392**, 936–941 (1998).
- Bading, H., Ginty, D. D. & Greenberg, M. E. Regulation of gene expression in hippocampal neurons by distinct calcium signaling pathways. *Science* **260**, 181–186 (1993).
- Deisseroth, K., Bito, H. & Tsien, R. W. Signalling from synapse to nucleus: postsynaptic CREB phosphorylation during multiple forms of hippocampal synaptic plasticity. *Neuron* **16**, 89–101 (1996).
- Bito, H., Deisseroth, K. & Tsien, R. W. CREB phosphorylation and dephosphorylation: a Ca^{2+} - and stimulus duration-dependent switch for hippocampal gene expression. *Cell* **87**, 1203–1214 (1996).
- Diesseroth, K., Heist, E. K. & Tsien, R. W. Translocation of calmodulin to the nucleus supports CREB phosphorylation in hippocampal neurons. *Nature* **392**, 198–202 (1998).
- Tao, X., Finkbeiner, S., Arnold, D. B., Shaywitz, A. J. & Greenberg, M. E. Ca^{2+} influx regulates BDNF transcription by a CREB family transcription factor-dependent mechanism. *Neuron* **20**, 709–726 (1998).
- Shieh, P. B., Hu, S.-C., Bobb, K., Timusk, T. & Ghosh, A. Identification of a signaling pathway involved in calcium regulation of BDNF expression. *Neuron* **20**, 727–740 (1998).
- Llina, R., Sugimori, M., Lin, J. W. & Cherksey, B. Blocking and isolation of a calcium channel from neurons in mammals and cephalopods utilizing a toxin fraction (FTX) from funnel-web spider poison. *Proc. Natl Acad. Sci. USA* **86**, 1689–1693 (1989).
- Mintz, I. M., Adams, M. E. & Bean, B. P. P-type calcium channels in rat central and peripheral neurons. *Neuron* **9**, 85–95 (1992).
- Bourinet, E. *et al.* Splicing of α_{1A} subunit gene generates phenotypic variants of P- and Q-type calcium channels. *Nature Neurosci.* **2**, 407–415 (1999).
- Betz, A., Okamoto, M., Benseler, F. & Brose, N. Direct interaction of the rat unc-13 homologue munc13-1 with the N-terminus of syntaxin. *J. Biol. Chem.* **272**, 2520–2526 (1997).
- Pevsner, J. *et al.* Specificity and regulation of a synaptic vesicle docking complex. *Neuron* **13**, 353–361 (1994).
- Fujita, Y. *et al.* Tomosyn: a syntaxin-1-binding protein that forms a novel complex in the neurotransmitter release process. *Neuron* **20**, 905–915 (1998).
- Sheng, Z.-H., Westenbroek, R. E. & Catterall, W. A. Physical link and functional coupling of presynaptic calcium channels and the synaptic vesicle docking/fusion machinery. *J. Bioenerg. Biomembr.* **30**, 335–345 (1998).

19. Sheng, Z.-H., Rettig, J., Takahashi, M. & Catterall, W. A. Identification of a syntaxin-binding site on N-type calcium channels. *Neuron* **13**, 1303–1313 (1994).
20. Bezprozvanny, I., Scheller, R. H. & Tsien, R. W. Functional impact of syntaxin on gating of N-type and Q-type calcium channels. *Nature* **378**, 623–626 (1995).
21. Martin-Moutot, N. *et al.* Interaction of SNARE complexes with P/Q-type calcium channels in rat cerebellar synaptosomes. *J. Biol. Chem.* **271**, 6567–6570 (1996).
22. Rettig, J. *et al.* Isoform-specific interaction of the α_{1A} subunits of brain Ca^{2+} channels with the presynaptic proteins syntaxin and SNAP-25. *Proc. Natl Acad. Sci. USA* **93**, 7363–7368 (1996).
23. Kim, D. K. & Catterall, W. A. Ca^{2+} -dependent and -independent interactions of the isoforms of the α_{1A} subunit of brain Ca^{2+} channels with presynaptic SNARE proteins. *Proc. Natl Acad. Sci. USA* **94**, 14782–14786 (1997).
24. Williamson, L. C., Halpern, J. L., Montecucco, C., Brown, J. E. & Neale, E. A. Clostridial neurotoxins and substrate proteolysis in intact neurons: botulinum neurotoxin C acts on synaptosomal-associated protein of 25 kDa. *J. Biol. Chem.* **271**, 7694–7699 (1996).
25. Honer, W. G., Hu, L. & Davies, P. Human synaptic proteins with a heterogeneous distribution in cerebellum and visual cortex. *Brain Res.* **609**, 9–20 (1993).
26. Daly, C. & Ziff, E. B. Post-transcriptional regulation of synaptic vesicle protein expression and the developmental control of synaptic vesicle formation. *J. Neurosci.* **17**, 2365–2375 (1997).
27. Gafni, J. *et al.* Xestospingins: potent membrane permeable blockers of the inositol 1,4,5-trisphosphate receptor. *Neuron* **19**, 723–733 (1997).
28. Gonzalez, G. A. & Montminy, M. R. Cyclic AMP stimulates somatostatin gene transcription by phosphorylation of CREB at serine 133. *Cell* **59**, 675–680 (1989).
29. Ginty, D. D., Bonni, A. & Greenberg, M. E. Nerve growth factor activates a Ras-dependent protein kinase that stimulates c-fos transcription via phosphorylation of CREB. *Cell* **77**, 713–725 (1994).
30. Carrion, A. M., Link, W. A., Ledo, F., Mellstrom, B. & Naranjo, J. R. DREAM is a Ca^{2+} -regulated transcriptional repressor. *Nature* **398**, 80–84 (1999).

Acknowledgements

We thank M. Gilbert, C. Santi, A. Stea and G. Zamponi for comments and discussions. We also thank W. Honer for providing the SP8 human syntaxin-1A monoclonal antibody and D. Brink and V. Leuranguer for tissue culture support. The research was supported by grants from the Medical Research Council (MRC) of Canada (T.P.S. and T.H.M.), postdoctoral support from the Amyotrophic Lateral Sclerosis Society of Canada (K.S.), and studentship and postdoctoral fellowship support from the MRC (H.G. and J.M., respectively). T.P.S. and T.H.M. are recipients of MRC of Canada Scientist awards.

Correspondence and requests for materials should be addressed to T.P.S. (e-mail: snutch@zoology.ubc.ca).

Extraintestinal dissemination of *Salmonella* by CD18-expressing phagocytes

Andrés Vazquez-Torres*, Jessica Jones-Carson†, Andreas J. Bäumlert‡, Stanley Falkow§, Raphael Valdivia||, William Brown*¶, Mysan Le*¶, Ruth Berggren*||, W. Tony Parks# & Ferric C. Fang*

* Departments of Medicine, Pathology, and Microbiology, University of Colorado Health Sciences Center, Denver, Colorado 80262, USA

† Department of Immunology, National Jewish Center for Immunology and Respiratory Medicine, Denver, Colorado 80262, USA

‡ Department of Medical Microbiology and Immunology, Texas A&M University, College Station, Texas 77843-1114, USA

§ Department of Microbiology and Immunology, Stanford University School of Medicine, Stanford, California 94305, USA

¶ Department of Medicine, Veterans Affairs Medical Center, Denver, Colorado 80262, USA

Laboratory of Cell Regulation and Carcinogenesis, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892, USA

Specialized epithelia known as M cells overlying the lymphoid follicles of Peyer's patches are important in the mucosal immune system, but also provide a portal of entry for pathogens such as *Salmonella typhimurium*, *Mycobacterium bovis*, *Shigella flexneri*, *Yersinia enterocolitica* and reoviruses^{1–4}. Penetration of intestinal

M cells and epithelial cells by *Salmonella typhimurium* requires the invasion genes of *Salmonella* Pathogenicity Island 1 (SPI1)^{3,5–9}. SPI1-deficient *S. typhimurium* strains gain access to the spleen following oral administration and cause lethal infection in mice⁵ without invading M cells^{3,9} or localizing in Peyer's patches¹⁰, which indicates that *Salmonella* uses an alternative strategy to disseminate from the gastrointestinal tract. Here we report that *Salmonella* is transported from the gastrointestinal tract to the bloodstream by CD18-expressing phagocytes, and that CD18-deficient mice are resistant to dissemination of *Salmonella* to the liver and spleen after oral administration. This CD18-dependent pathway of extraintestinal dissemination may be important for the development of systemic immunity to gastrointestinal pathogens, because oral challenge with SPI1-deficient *S. typhimurium* elicits a specific systemic IgG humoral immune response, despite an inability to stimulate production of specific mucosal IgA.

As phagocytic cells transmigrate across normal tissue barriers¹¹ and *Salmonella* can survive within mononuclear phagocytic cells¹², we proposed that tissue macrophages transport *Salmonella* from the intestine to the systemic circulation; CD18-deficient mice¹³ provided a model in which to test this hypothesis. Among its many functions, the β 2-integrin CD18 mediates leukocyte transmigration in response to various stimuli, including *S. typhimurium*^{14,15}.

We introduced mutations in the *invA* gene into wild-type *S. typhimurium* ATCC 14028s to inactivate the SPI1 invasion genes, and added a mutation in the *lpfC* fimbrial gene¹⁶ to impair targeting of Peyer's patches further. An *aroA* mutation in the gene encoding 3-enol-pyruvylshikimate synthase was incorporated to reduce virulence and allow the survival of infected animals. CD18-deficient mice infected orally with the *invA*–*lpfC*–*aroA* mutant *S. typhimurium* strain harboured 100-fold fewer bacteria in the spleen and liver than did similarly infected congenic C57BL/6 controls (Fig. 1a). The resistance of CD18-deficient mice to *Salmonella* dissemination from the gastrointestinal tract cannot be explained by enhanced bacterial clearance from systemic sites, because the CD18-deficient mouse strain was found to be more susceptible than congenic C57BL/6 mice to splenic and hepatic infection following intraperitoneal challenge with *invA*–*lpfC*–*aroA* mutant *S. typhimurium* (Fig. 1b). Moreover, peritoneal macrophages from CD18-deficient mice were less efficient at killing *Salmonella* *in vitro* than control macrophages from C57BL/6 mice (data not shown). These observations indicate that a CD18-dependent host process may enable *S. typhimurium* to

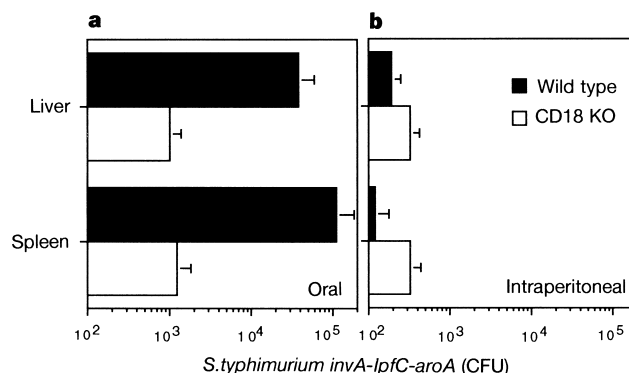


Figure 1 CD18-dependent dissemination of invasion-deficient *S. typhimurium* from the gastrointestinal tract to systemic organs. CD18 β 2 integrin-deficient mice are more resistant than congenic C57BL/6 controls to systemic dissemination of *invA*–*lpfC*–*aroA* mutant bacteria following oral (a) but not intraperitoneal (b) inoculation. The spleens of C57BL/6 mice contain 100-fold more bacteria than those from CD18-deficient controls after oral administration, but 2–3-fold fewer bacteria after intraperitoneal inoculation. The data represent the mean $\pm$ s.e.m. of 5–10 mice from two independent experiments.

|| Present addresses: Department of Biochemistry, University of California at Berkeley, Berkeley, California 94720, USA (R.V.); Department of Medicine, University of Texas Southwestern Medical Center, Dallas, Texas 75325, USA (R.B.).

disseminate systematically without invading M cells and colonizing Peyer's patches.

S. typhimurium carrying mutations in *invA*, *lpfC* and *aroA* were found in the blood stream of C57BL/6 mice as early as 15 min after oral inoculation. Rapid dissemination was observed whether the bacteria were fed orally or were introduced directly into the stomach by cannula, indicating that dissemination may occur in the gastrointestinal tract and not the oral mucosa. In concordance with studies of *Salmonella typhi*¹⁷, the level of bacteraemia peaked at roughly 2×10^3 colony-forming units per millilitre of blood 30 min after oral inoculation and declined over the next several hours. All bacteria recovered from the blood were biochemically confirmed to be *S. typhimurium*. Despite high numbers of microorganisms present in the spleen and liver (Fig. 1a), the bacteremia completely resolved by 2–5 days after inoculation. The rapid time course and absence of histopathological abnormalities throughout the intestine (data not shown) indicated that gastrointestinal

inflammation is not required for *Salmonella* invasion of the blood-stream.

An *invA*–*aroA* mutant *S. typhimurium* strain stably expressing green fluorescent protein (GFP) was constructed to determine whether blood-borne bacteria were localized within cells. Flow cytometric analysis of peripheral blood collected from C57BL/6 mice 1 h after oral challenge with the GFP-tagged invasion-deficient *S. typhimurium* strain showed that fluorescent bacteria were exclusively within CD18-expressing leukocytes (Fig. 2a–c). All of these CD18 positive leukocytes co-expressed the CD11a α and CD11b α subunits, forming the LFA-1 and Mac-1 complexes, respectively. In addition, roughly 20% of the GFP-containing cells co-expressed the CD18–CD11c heterodimer. Giemsa staining of flow cytometrically sorted CD18-expressing leukocytes harbouring GFP-positive bacteria allowed visualization of *S. typhimurium* within mononuclear phagocytes (Fig. 2d, e), which was further confirmed by fluorescence microscopy (Fig. 2f). The α -integrin expression and mor-

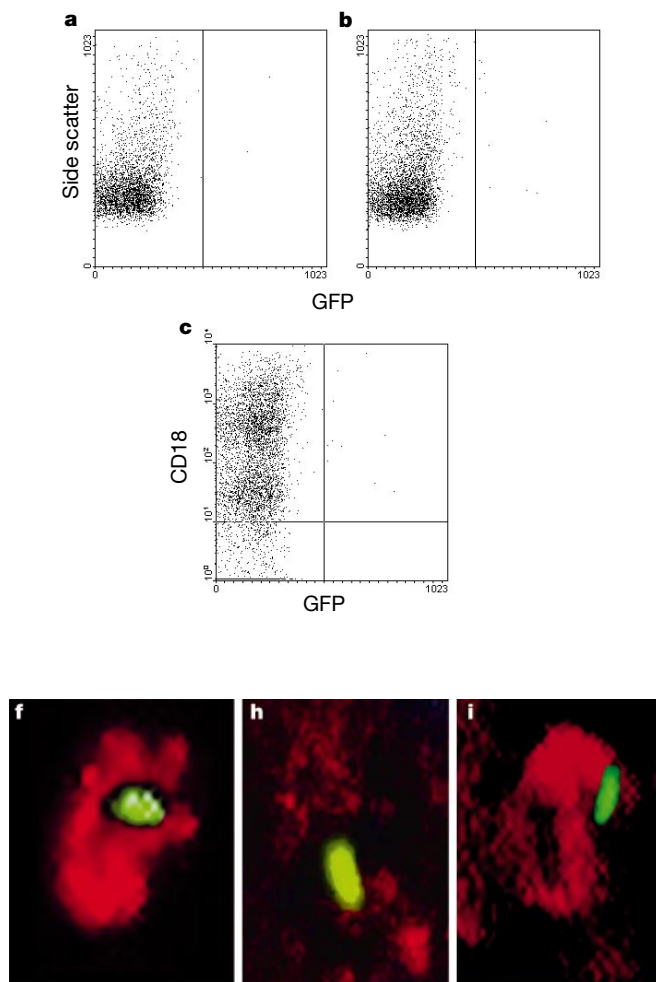


Figure 2 CD18-expressing cells of monocyte–macrophage lineage mediate extra-intestinal dissemination of invasion-deficient *S. typhimurium*. *S. typhimurium* carrying mutations in *invA*, *lpfC* and *aroC* are detectable in the bloodstream of C57BL/6 mice within minutes after oral inoculation. In contrast to white blood cells from uninfected controls (not shown) or C57BL/6 mice infected with GFP-negative invasion-deficient *S. typhimurium* strain AF982 (**a**), roughly 0.09% of the white blood cells isolated from C57BL/6 mice were found to carry GFP-tagged invasion-deficient *S. typhimurium* one hour after oral challenge (**b**) associated with a small population of CD18-expressing host cells (**c**). Roughly 18 ± 4 CD18-positive cells carrying GFP-expressing *S. typhimurium* AF981 were detected per 2×10^4 peripheral-blood leukocytes analysed in orally infected C57BL/6 mice, whereas parallel studies using GFP-negative *S. typhimurium* AF982 detected only 2 ± 1 double-positive cells per 2×10^4 peripheral-blood leukocytes,

reflecting baseline autofluorescence. Flow cytometric sorting of CD18⁺ GFP⁺ cells showed bacteria residing within cells of monocyte–macrophage lineage (**d**, **e**) (original magnification, 1,800 $\times$). **f**, A green GFP-expressing bacterium associated with a red phycoerythrin-labelled CD18-expressing peripheral-blood leukocyte (original magnification, 8,000 $\times$). **g**, CD18-expressing monocytes isolated from a noninfected control (original magnification, 1,800 $\times$). **h**, **i**, CD18-expressing leukocytes harbouring GFP-positive invasion-deficient *S. typhimurium* were visualized by confocal microscopy in the lamina propria of the ileal villi and by fluorescence microscopy in blood as early as 5 min after inoculation of a ligated intestinal loop (magnification, 8,000 $\times$). Arrows indicate the location of intracellular bacteria. The data shown were obtained in 12 independent experiments.

phology of bacteria-containing cells relative to control CD18-expressing monocytes (Fig. 2g) indicate that the bacteria-containing cells may belong to the monocyte–macrophage lineage, possibly consisting of dendritic cells and/or tissue macrophages, although migration of macrophages from peripheral sites to the bloodstream is not conventionally believed to occur. Immunohistochemical studies have shown CD18-expressing mononuclear cells in

normal intestinal mucosa¹⁸, which presumably represents the site of origin of the *Salmonella*-containing cells. Bloodstream trafficking of *Salmonella* within CD18-expressing phagocytes may help to account for the striking propensity of *Salmonella* to infect tissues such as aortic atherosclerotic plaques¹⁹ that are typically infiltrated with mononuclear cells.

Inoculation of invasion-deficient *S. typhimurium* into a ligated

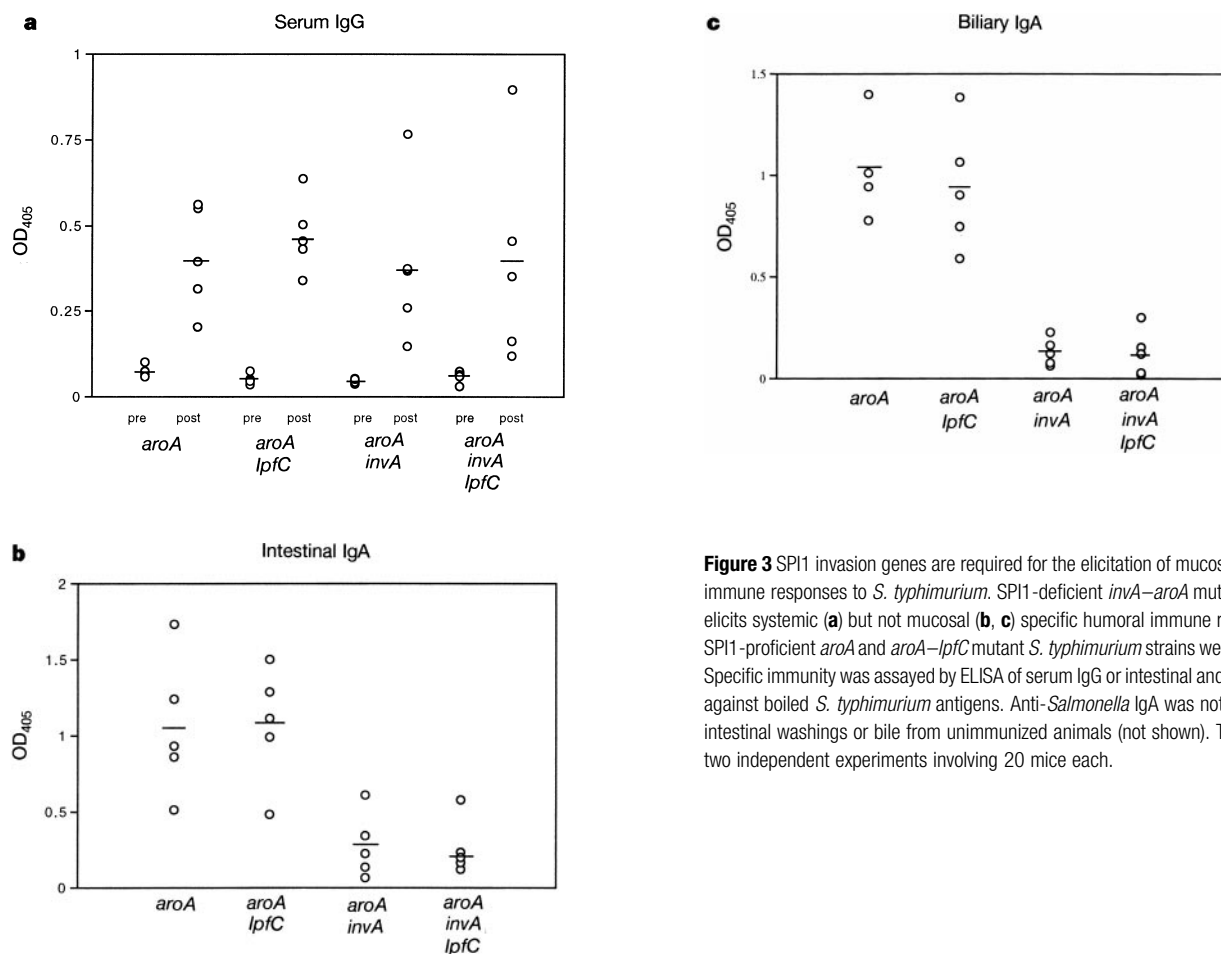


Figure 3 SPI1 invasion genes are required for the elicitation of mucosal but not systemic immune responses to *S. typhimurium*. SPI1-deficient *invA*–*aroA* mutant *S. typhimurium* elicits systemic (a) but not mucosal (b, c) specific humoral immune responses. Isogenic SPI1-proficient *aroA* and *aroA*–*lpfC* mutant *S. typhimurium* strains were used as controls. Specific immunity was assayed by ELISA of serum IgG or intestinal and biliary IgA directed against boiled *S. typhimurium* antigens. Anti-*Salmonella* IgA was not detectable in intestinal washings or bile from unimmunized animals (not shown). The data represent two independent experiments involving 20 mice each.

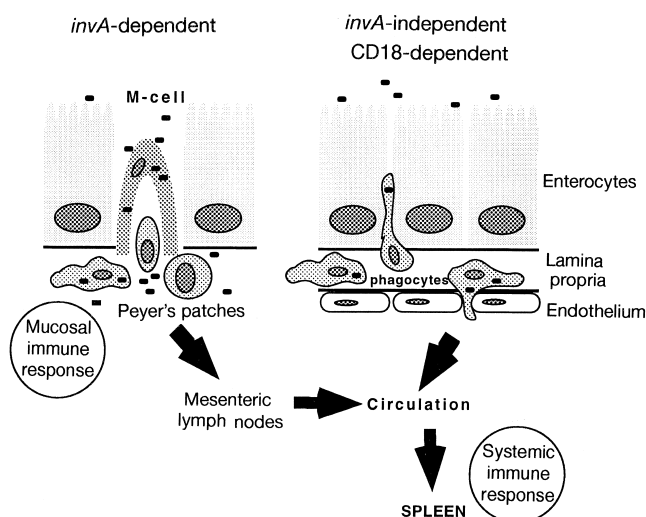


Figure 4 Model for functional and anatomical compartmentalization of the immune response to *Salmonella*. The *invA*-dependent SPI1 invasion genes allow *S. typhimurium* to enter M cells and gain access to gut-associated lymphoid tissue, where stimulation of a specific mucosal immune response occurs. Alternatively, gastrointestinal CD18-expressing cells of monocyte–macrophage lineage can take up *Salmonella* in a SPI1-independent process and transport the bacteria through the bloodstream to the spleen, resulting in the elicitation of a specific systemic immune response.

intestinal loop showed that CD18-dependent extraintestinal dissemination can originate in the ileum. CD18-expressing cells harbouring GFP-positive, invasion-deficient *Salmonella* AF981 could be observed in transit across the lamina propria of the ileal villi (Fig. 2h) and in the bloodstream (Fig. 2i) as early as 5 min after intraluminal inoculation.

Trans epithelial antigen transport by M cells exposes subadjacent follicular leukocytes to luminal antigens, initiating a specific mucosal immune response^{4,20}. Indeed, a noninvasive SPII-deficient *S. typhimurium* strain failed to elicit an intestinal or biliary IgA response, in contrast to isogenic invasion-proficient controls (Fig. 3b, c). Nevertheless, noninvasive bacteria retain the ability to elicit a systemic immune response as measured by the presence of *Salmonella*-specific IgG in serum (Fig. 3a). Trafficking CD18-positive phagocytes may facilitate antigen presentation in the spleen, while simultaneously providing a route of extraintestinal dissemination for pathogenic microorganisms capable of intracellular survival. A requirement for CD18-positive cells in translocation of *Y. enterocolitica* to the spleen has been reported²¹. *Y. enterocolitica* was proposed to disseminate to the spleen only after invading Peyer's patches, whereas our work indicates that extraintestinal bacterial dissemination can occur without colonization of gut-associated lymphoid tissue.

Our observations indicate that there may be a functional and anatomic compartmentalization of the immune system dependent upon the route of bacterial uptake (Fig. 4). The host appears to use CD18-expressing cells of monocyte-macrophage lineage to survey and initiate systemic immune responses to gastrointestinal microbial antigens, whereas M cells control the independent generation of mucosal immune responses. In addition to its well characterized ability to invade M cells and gain access to gut-associated lymphoid tissue, *Salmonella* can disseminate from the gastrointestinal tract by CD18-expressing phagocytic cells. The CD18-dependent immune surveillance mechanism appears to facilitate the development of a systematic immune response, but the trade-off is a pathway of entry for pathogens adapted to survival within phagocytic cells and a potential route for spread of these pathogens throughout the body. □

Methods

Bacterial strains

S. typhimurium derivatives of wild-type ATCC 14028¹² carrying mutations in *invA* (*invA::TnphoA*) and *lpfC* (*lpfC::cat*) have been described¹⁰. These mutations were combined with *aroA::Tn10* from strain CL1509 (ref. 22) by P22-mediated transduction to yield *S. typhimurium* strains AJB74 (*lpfC-aroA* mutant), AJB82 (*invA-aroA* mutant) and FF728 (*invA-lpfC-aroA* mutant). An *invA* mutant *S. typhimurium* strain stably expressing *gfp* was constructed to allow visualization of blood-borne bacteria. An *rpsM::gfp* fusion²³ was inserted into pGP704K (*oriR6K-mob-aph*) and the resulting plasmid conjugated into wild-type *S. typhimurium* SL1344 to create strain SM022 (*rpsM::gfp*) (a gift from S. Moller). Expression of *gfp* was confirmed by fluorescence microscopy and flow cytometry. The *invA* gene sequence of *S. typhimurium* was amplified by PCR²⁴, cloned into plasmid pCRII (Invitrogen), moved into suicide vector pEP185.2 (ref. 25) and transformed into *Escherichia coli* strain S17-1 λ pir²⁶. This strain was conjugated with *S. typhimurium* IR715 (ref. 27) to yield AJB634 (*invA::pEP185.2*). Interruption of *invA* was confirmed genotypically by Southern blot and phenotypically by competitive infection experiments in mice (data not shown). The *invA::pEP185.2*, *aroA::tet* and *rpsM::aph* mutations were transduced into *S. typhimurium* strain 14028 to yield *S. typhimurium* AF981 (*invA-aroA-gfp*⁺) and AF982 (*invA-aroA*).

Detection of extraintestinal bacterial dissemination

Dissemination of *S. typhimurium* FF728 (*invA-lpfC-aroA* mutant) was determined as described¹⁰. C57BL/6 or congenic CD18-deficient mice¹³ (Jackson Laboratories) were fasted for 3 h before oral administration of 5×10^8 colony-forming units of stationary phase *S. typhimurium* FF728, a standard dose for *aroA*-mutant *S. typhimurium* vaccines²⁸. Additional mice were inoculated intraperitoneally with about 100 colony-forming units (c.f.u.) of *S. typhimurium* FF728. Seven days after inoculation, spleens and livers were homogenized for quantification of bacterial burden, or fixed in 10% formalin for haematoxylin-eosin staining.

Detection of bacteraemia

C57BL/6 mice were anesthetized and peripheral blood was collected in heparinized tubes

at several time points after oral or intraluminal bacterial challenge. Host cells were lysed with 0.5% sterile-filtered deoxycholic acid (Sigma) and blood-borne bacteria were quantified after culture on LB-agar plates. Bacteria were confirmed to be *S. typhimurium* FF728 and AF981 by subculture in triple sugar iron agar (Difco), expression of antibiotic markers and visualization by fluorescence microscopy.

Flow cytometry

S. typhimurium AF981 (*invA::pEP185.2 aroA::tet gfp*⁺) or AF982 (*invA::pEP185.2 aroA::tet*) were fed orally to C57BL/6 mice as above. One hour after infection, peripheral blood leukocytes isolated in a 44%/67% percoll gradient were first labelled with biotinylated anti-CD18, anti-CD11a, anti-CD11b and anti-CD11c monoclonal antibodies (PharMingen) or isotype controls, and then with phycoerythrin-conjugated avidin. The percentage of CD18 or CD11a/CD11b/CD11c (red) and GFP (green) double-positive leukocytes, gated by forward and side scatters, were analysed on a FACS Calibur flow cytometer. CD18 and GFP double-positive leukocytes were collected under sterile conditions using an Excel Coulter cell sorter and either cytocentrifuged onto a slide and stained by the Giemsa method (Sigma), or fixed in 3% paraformaldehyde and visualized by fluorescence microscopy.

Ligated intestinal loop model

C57BL/6 mice were fasted for 3–4 h before they were anesthetized by intraperitoneal inoculation of 1.5 mg pentobarbital sodium (Abbott Laboratories) per mouse³. The intestine was exposed through a small incision into the abdominal cavity, and a loop was constructed by ligating the ileum at the ileocecal junction and at a site 3–4 cm proximal to the cecum. Roughly 5×10^8 c.f.u. of *S. typhimurium* strain AF981 (*invA::pEP185.2 aroA::tet gfp*⁺) were inoculated into the loop. Peripheral blood leukocytes found 5–60 min after inoculation were isolated on a percoll gradient, labelled with an anti-CD18 monoclonal antibody, fixed in 3% paraformaldehyde and visualized by immunofluorescence microscopy as described above. In selected mice, ileal mucosal tissue was embedded in optimal cutting temperature (Sakura Finetek USA Inc.), sliced in frozen sections, stained with an anti-CD18 monoclonal antibody and processed for confocal microscopy.

Measurement of humoral immune responses

Groups of BALB/c mice were pretreated with 10% NaHCO₃ and inoculated orally with 2×10^8 c.f.u. of *S. typhimurium* strain CL1509 (*aroA* mutant), AJB74 (*lpfC-aroA* mutant), AJB82 (*invA-aroA* mutant) or FF728 (*invA-lpfC-aroA* mutant) at intervals of 14 days during a period of 28 days. Serum was collected 1 day before each immunization and at the time of killing. Two weeks after the final boost, intestinal washings were collected in PBS containing 50 mM EDTA, 0.1 mg ml⁻¹ soybean trypsin inhibitor and 1% (v/v) phenylmethylsulphonyl fluoride. Serum diluted 1:100 and intestinal washings or bile diluted 1:10 in PBS with 0.3% gelatin were assayed for *Salmonella*-specific IgG and IgA by enzyme-linked immunosorbent assay (ELISA) as described²⁹. Boiled *S. typhimurium* cells were used as capture antigen, and horseradish peroxidase-labelled rabbit anti-mouse IgG or IgA (Zymed) were used as conjugates. The results are expressed as the average OD₄₀₅ from duplicate wells representing individual mice.

Received 16 July; accepted 18 August 1999.

1. Wolf, J. L. *et al.* Intestinal M cells: a pathway for entry of reovirus into the host. *Science* **212**, 471–472 (1981).
2. Wassef, J. S., Keren, D. F. & Mailloux, J. L. Role of M cells in initial antigen uptake and in ulcer formation in the rabbit intestinal loop model of shigellosis. *Infect. Immun.* **57**, 858–863 (1989).
3. Jones, B. D., Ghori, N. & Falkow, S. *Salmonella typhimurium* initiates murine infection by penetrating and destroying the specialized epithelial M cells of the Peyer's patches. *J. Exp. Med.* **180**, 15–23 (1994).
4. Neutra, M. R., Frey, A. & Kraehenguhl, J. P. Epithelial M cells: gateways for mucosal infection and immunization. *Cell* **86**, 345–348 (1996).
5. Galan, J. E. & Curtiss, R. III. Cloning and molecular characterization of genes whose products allow *Salmonella typhimurium* to penetrate tissue culture cells. *Proc. Natl Acad. Sci. USA* **86**, 6383–6387 (1989).
6. Lee, C. A., Jones, B. D. & Falkow, S. Identification of a *Salmonella typhimurium* invasion locus by selection for hyperinvasive mutants. *Proc. Natl Acad. Sci. USA* **89**, 1847–1851 (1992).
7. Francis, C. L., Ryan, T. A., Jones, B. D., Smith, S. J. & Falkow, S. Ruffles induced by *Salmonella* and other stimuli direct macropinocytosis of bacteria. *Nature* **364**, 639–642 (1993).
8. Collazo, C. M. & Galan, J. E. The invasion-associated type III protein secretion system in *Salmonella*—a review. *Gene* **192**, 51–59 (1997).
9. Pehheiter, K. L., Mathur, N., Giles, D., Gahlen, T. & Jones, B. D. Non-invasive *Salmonella typhimurium* mutants are avirulent because of an inability to enter and destroy M cells of ileal Peyer's patches. *Mol. Microbiol.* **24**, 697–709 (1997).
10. Baumler, A. J., Tsolis, R. M., Valentine, P. J., Ficht, T. A. & Heffron, F. Synergistic effect of mutations in *invA* and *lpfC* on the ability of *Salmonella typhimurium* to cause murine typhoid. *Infect. Immun.* **65**, 2254–2259 (1997).
11. Shang, X. Z. & Issekutz, A. C. Contribution of CD11a/CD18, CD11b/CD18, ICAM-1 (CD54) and -2 (CD102) to human monocyte migration through endothelium and connective tissue fibroblast barriers. *Eur. J. Immunol.* **28**, 1070–1079 (1998).
12. Fields, P. L., Swanson, R. V., Haidaris, T. G. & Heffron, F. Mutants of *Salmonella typhimurium* that cannot survive within the macrophage are avirulent. *Proc. Natl Acad. Sci. USA* **83**, 5189–5193 (1986).
13. Wilson, R. W. *et al.* Gene targeting yields a CD18-mutant mouse for study of inflammation. *J. Immunol.* **151**, 1571–1578 (1993).
14. Kishimoto, T. K. & Anderson, D. C. In *Inflammation: Basic Principles and Clinical Correlates* 2nd edn (eds Gallin, J. I., Goldstein, I. M. & Snyderman, R.) 353–406 (Raven, New York, 1992).
15. Conlan, J. W. & North, R. J. *Listeria monocytogenes*, but not *Salmonella typhimurium*, elicits a CD18-

- independent mechanism of neutrophil extravasation into the murine peritoneal cavity. *Infect. Immun.* **62**, 2702–2706 (1994).
16. Bauml, A. J., Tsolis, R. M. & Heffron, F. The *lpf* fimbrial operon mediates adhesion to murine Peyer's patches. *Proc. Natl Acad. Sci. USA* **93**, 279–283 (1996).
 17. Gerichter, C. B. The dissemination of *Salmonella typhi*, *S. paratyphi A* and *S. paratyphi B* through the organs of the white mouse by oral infection. *J. Hyg.* **58**, 307–319 (1960).
 18. Bernstein, C. N., Sargent, M., Gallatin, W. M. & Wilkins, J. β_2 -integrin/intercellular adhesion molecule (ICAM) expression in the normal human intestine. *Clin. Exp. Immunol.* **106**, 160–169 (1996).
 19. Oskoui, R., Davis, W. & Gomes, M. N. *Salmonella aortitis*. A report of a successfully treated case with a comprehensive review of the literature. *Arch. Int. Med.* **153**, 517–525 (1993).
 20. Neutra, M. R. Current concepts in mucosal immunity. V. Role of M cells in transepithelial transport of antigens and pathogens to the mucosal immune system. *Am. J. Physiol.* **274**, G785–791 (1998).
 21. Autenrieth, I. B., Kempf, V., Sprinz, T., Preger, S. & Schnell, A. Defense mechanisms in Peyer's patches and mesenteric lymph nodes against *Yersinia enterocolitica* involve integrins and cytokines. *Infect. Immun.* **64**, 1357–1368 (1996).
 22. Tsolis, R. M., Bauml, A. J., Stojiljkovic, I. & Heffron, F. Fur regulon of *Salmonella typhimurium*: identification of new iron-regulated genes. *J. Bacteriol.* **177**, 4628–4637 (1995).
 23. Valdivia, R. H. & Falkow, S. Bacterial genetics by flow cytometry: rapid isolation of *Salmonella typhimurium* acid-inducible promoters by differential fluorescence induction. *Mol. Microbiol.* **22**, 367–378 (1996).
 24. Rahn, K. et al. Amplification of an *invA* gene sequence of *Salmonella typhimurium* by polymerase chain reaction as a specific method of detection of *Salmonella*. *Mol. Cell. Probes* **6**, 271–279 (1992).
 25. Kinder, S. A., Badger, J. L., Bryant, G. O., Pepe, J. C. & Miller, V. L. Cloning of the *YenI* restriction endonuclease and methyltransferase from *Yersinia enterocolitica* serotype O:8 and construction of a transformable R-M+ mutant. *Gene* **136**, 271–275 (1993).
 26. Miller, V. L. & Mekalanos, J. J. A novel suicide vector and its use in construction of insertion mutations; osmoregulation of outer membrane proteins and virulence determinants requires *toxR*. *J. Bacteriol.* **170**, 2575–2583 (1988).
 27. Stojiljkovic, I., Bauml, A. J. & Heffron, F. Ethanolamine utilization operon of *Salmonella typhimurium*: nucleotide sequence, protein expression and mutational analysis of the *cbhB* *eutE* *eutG* *eutH* gene cluster. *J. Bacteriol.* **177**, 1357–1366 (1995).
 28. Valentine, P. J., Devore, B. P. & Heffron, F. Identification of three highly attenuated *Salmonella typhimurium* mutants that are more immunogenic and protective in mice than a prototypical *aroA* mutant. *Infect. Immun.* **66**, 3378–3383 (1998).
 29. Berggren, R. E. et al. HIV gp120-specific cell-mediated immune responses in mice after oral immunization with recombinant *Salmonella*. *J. AIDS* **10**, 489–495 (1995).

Acknowledgements

This work was supported in part by grants from the NIH (A.V.T., J.J.C., A.B., R.B. and F.F.), the United States Department of Agriculture (A.B.), and the James Biundo Foundation (F.F.). The authors are grateful to E. Balish for constructive review of the manuscript and thank F. Heffron and S. Moller for the gifts of strains CL1509 and SM022, respectively.

Correspondence and requests for materials should be addressed to F.C.F. (e-mail: ferric.fang@uchsc.edu).

Integrin cytoplasmic tyrosine motif is required for outside-in α IIB β 3 signalling and platelet function

Debbie A. Law*, Francis R. DeGuzman*, Patrick Heiser*, Kathleen Ministri-Madrid*, Nigel Killeen† & David R. Phillips*

* COR Therapeutics, Inc., 256 East Grand Avenue, South San Francisco, California 94080, USA

† Department of Microbiology and Immunology, University of California, San Francisco, California 94143-0414, USA

Integrins not only bind adhesive ligands¹, they also act as signalling receptors². Both functions allow the integrin α IIB β 3 to mediate platelet aggregation³. Platelet agonists activate α IIB β 3 (inside-out signalling) to allow the binding of soluble fibrinogen. Subsequent platelet aggregation leads to outside-in α IIB β 3 signalling, which results in calcium mobilization⁴, tyrosine phosphorylation of numerous proteins^{5,6} including β 3 itself⁷, increased cytoskeletal reorganisation⁸ and further activation of α IIB β 3 (ref. 2). Thus, outside-in signals enhance aggregation, although the mechanisms and functional consequences of specific

signalling events remain unclear. Here we describe a mouse that expresses an α IIB β 3 in which the tyrosines in the integrin cytoplasmic tyrosine motif have been mutated to phenylalanines. These mice are selectively impaired in outside-in α IIB β 3 signalling, with defective aggregation and clot-retraction responses *in vitro*, and an *in vivo* bleeding defect which is characterized by a pronounced tendency to rebleed. These data provide evidence for an important role of outside-in signalling in platelet physiology. Furthermore, they identify the integrin cytoplasmic tyrosine motif as a key mediator of β -integrin signals and a potential target for new therapeutic agents.

The β 3 subunit of α IIB β 3 contains two cytoplasmic tyrosine residues and is phosphorylated upon platelet aggregation⁷. The tyrosines form part of a motif common to several β -integrin subunits, which we refer to as the integrin cytoplasmic tyrosine (ICY) motif, consisting of two tyrosines separated by 11–19 residues with the upstream tyrosine in the context of NPXY and the downstream tyrosine in NxxY, both potential phosphotyrosine binding (PTB) recognition sites⁹. For α IIB β 3, biochemical studies suggest that both tyrosines are phosphorylated and are involved in the recruitment of signalling (SHC and GRB2)⁷ and cytoskeletal (myosin) proteins¹⁰. To test the importance of the ICY motif and β 3 tyrosine phosphorylation in platelet physiology, we generated mice with an ICY-mutant β 3 gene carrying phenylalanines in place of two tyrosines (Y747 and Y759). Phenylalanines substitute for tyrosine with minimal effects on protein structure¹¹, and interactions of β -integrin subunits with other proteins that are not dependent on tyrosine phosphorylation can be maintained after tyrosine-to-phenylalanine substitutions^{12,13}. Mice homozygous (diYF) and heterozygous for the mutant β 3 gene were viable, fertile and born at the expected Mendelian ratio. The homozygous mutant mice had normal platelet counts and expressed wild-type levels of α IIB β 3 (Fig. 1).

In contrast to wild-type platelets, diYF platelets failed to aggregate when stimulated with 0.05 units ml⁻¹ thrombin (Fig. 2a). At

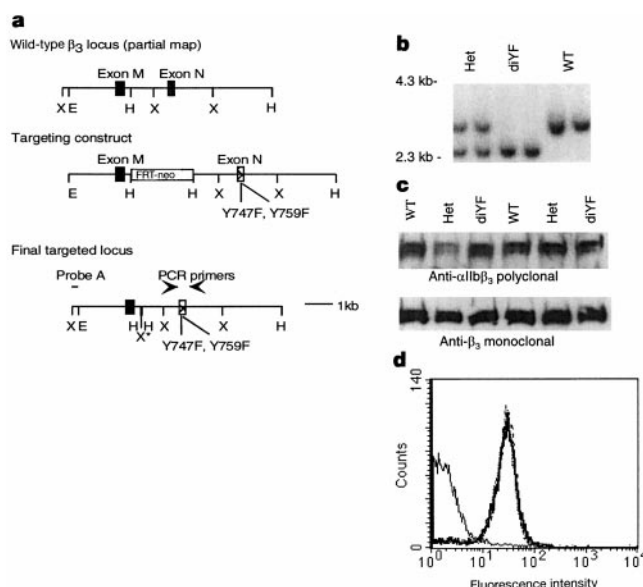


Figure 1 Generation of β 3 diYF mice. **a**, Targeting vectors. X, *XbaI*; E, *EcoRI*; H, *HindIII*; X*, FRT recognition site remnant with *XbaI*. **b**, Southern analysis of tail DNA digested with *XbaI*. Wild-type (WT) allele, 3.0 kb; mutant allele, 2.4 kb using Probe A. **c**, Western blot analysis of β 3 protein levels in platelet lysates from WT, heterozygous (Het) and diYF mutant mice. **d**, Flow cytometric analysis of β 3 levels on platelets. Geometric mean fluorescence $\pm$ s.d.: WT, 24.4 $\pm$ 2.9; heterozygote, 23.7 $\pm$ 3.4, and mutant, 22.7 $\pm$ 3.8 (n = 8). Negative control antibody plot is shown on the left.

0.1 units ml^{-1} thrombin, diYF platelets formed aggregates, but these were markedly unstable and disaggregated into individual platelets. At high doses of thrombin (≥ 1 units ml^{-1}), however, the aggregation reactions of wild-type, heterozygous and diYF platelets were indistinguishable. The same transient aggregation phenotype was observed in response to the weaker platelet agonist, ADP (Fig. 2b), although spontaneous disaggregation of the diYF platelets occurred over a broad concentration range (up to 100 μM ADP). As with thrombin, heterozygous platelets consistently showed an intermediate phenotype. Aggregation of wild-type platelets caused tyrosine phosphorylation of $\beta 3$, a reaction that did not occur in diYF platelets (data not shown). The data indicate that tyrosine phosphorylation of the $\beta 3$ ICY motif is required for optimal and stable aggregation in response to a weak agonist, but that high concentrations of thrombin, a more potent agonist, will overcome the signalling defect caused by the ICY mutation.

The initial slopes of the aggregometer tracings when fibrinogen was present in the platelet-rich plasma (ADP-induced aggregations) were similar for all three genotypes (Fig. 2b), which indicated that the initial activation and extent of fibrinogen binding (events dependent on inside-out $\alpha\text{IIb}\beta 3$ signalling²) were comparable. Consistent with this, stimulated platelets from diYF and wild-type mice bound similar levels of fluorescein-labelled soluble fibrinogen (Fig. 2c). In other experiments, the diYF mutation also did not impact on the fibrinogen content of platelets (data not shown). Thus, the ICY mutation does not block $\alpha\text{IIb}\beta 3$ inside-out signalling, but instead selectively abrogates outside-in signalling.

Clot retraction appears to be an important part of thrombus consolidation and is $\alpha\text{IIb}\beta 3$ dependent¹⁴. diYF platelets were defective in retracting clots; this was most notable when 1 unit ml^{-1} of thrombin was used to induce clot formation (Fig. 3a), but a consistent, albeit smaller, difference was noted at higher thrombin concentrations (Fig. 3b). These data substantiate conclusions from cell transfection studies^{10,15}, and again establish a key role for ICY-mediated outside-in $\alpha\text{IIb}\beta 3$ signalling in platelet function.

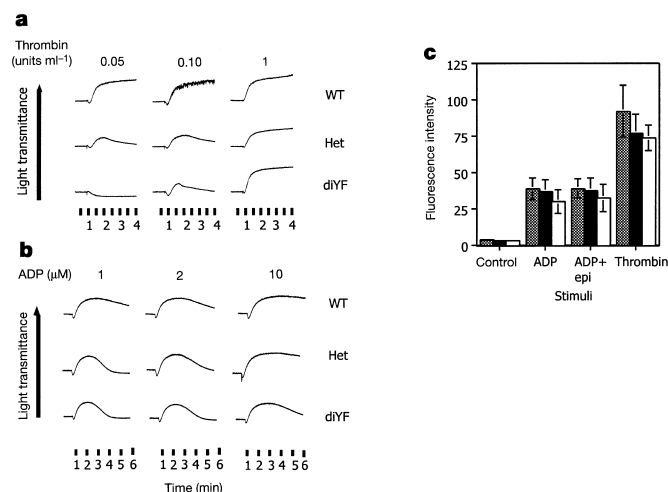


Figure 2 Aggregation and soluble fibrinogen binding in murine platelets. **a**, Thrombin-induced aggregation of washed platelets. Wild-type platelets failed to respond at 0.03 units ml^{-1} (1 unit of thrombin = 9 nM). **b**, ADP-induced aggregation of platelets in platelet-rich plasma. **a**, **b**, Data are representative of three experiments. **c**, Flow cytometric analysis of fluorescein isothiocyanate (FITC)-conjugated fibrinogen binding to platelets stimulated with ADP (50 μM), ADP (10 μM) and epinephrine (epi, 25 μM), or thrombin (0.1 units ml^{-1}), or unstimulated (control). Duplicate samples were analysed. Graph shows compilation of geometric mean $\pm$ s.d. from WT (grey; $n = 11$), heterozygotes (black; $n = 9$) and diYF (white; $n = 10$).

As a significant proportion of $\beta 3$ redistributes to the Triton-X-100 (TX-100)-insoluble cytoskeletal fraction following platelet aggregation¹⁶, a redistribution that favours tyrosine-phosphorylated $\beta 3$ (ref. 10), we wondered whether the diYF mutation might affect stable aggregation and clot retraction by disrupting the interaction of phospho- $\beta 3$ with the cytoskeleton. However, there was no difference in the ability of wild-type and mutant $\beta 3$ to redistribute to the TX-100 detergent-insoluble cytoskeletal fraction upon platelet aggregation (data not shown). Thus, although it is possible that the loss of myosin binding caused by the diYF mutation may impact platelet function, it does not appear to be caused by a change in interaction with the cytoskeleton.

As a direct test of the physiological significance of outside-in $\alpha\text{IIb}\beta 3$ signalling *in vivo*, we examined the impact of the ICY mutation on haemostasis using a transected tail cut model of bleeding time. As shown in Fig. 4, diYF mice showed a pronounced tendency to rebleed after transient haemostasis ($P < 0.001$, 95% confidence level for mutant compared with wild-type). This phenotype is consistent with our *in vitro* data and indicative of instability in haemostatic-plug formation *in vivo* being caused by the loss of ICY-dependent $\alpha\text{IIb}\beta 3$ signalling. Whereas rebleeding was clearly affected by the ICY mutation, there was no firm evidence for a prolongation of bleeding time when diYF animals were compared with wild-type littermates ($P = 0.232$, Mann-Whitney non-parametric test). A large number of progeny from heterozygote crossings were used for these studies (wild-type, $n = 42$; heterozygotes, $n = 72$; diYF, $n = 32$); however, the majority of wild-type animals had bleeding times that fell within a 108.5-s range (first

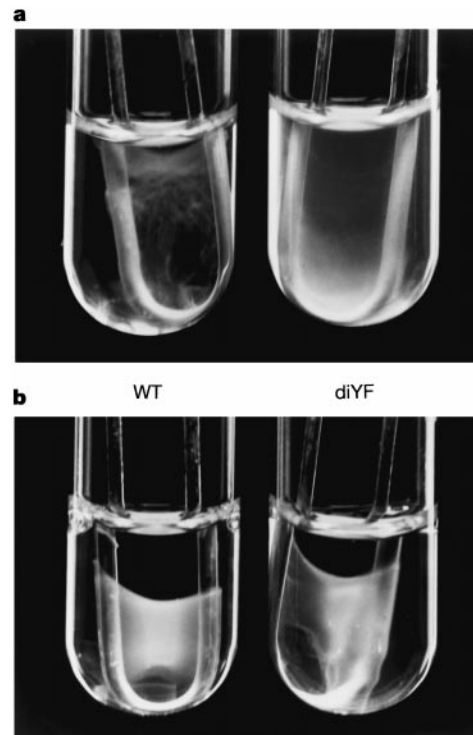


Figure 3 Retraction of fibrin clots. Photographs show the degree of retraction after 2 h in WT and diYF PRP samples treated with 1 units ml^{-1} (9 nM) thrombin (**a**), 10 units ml^{-1} (90 nM) thrombin (**b**), with similar results being observed in two more experiments. diYF mice had a 24–46% defect in their abilities to retract clots when compared with WT (when thrombin was used at 3 units ml^{-1} ($n = 5$) or 10 units ml^{-1} ($n = 3$)). At the 1 units ml^{-1} thrombin dose, visual inspection indicated a roughly 50–75% defect in the clot-retraction capabilities of platelets from diYF compared with WT mice ($n = 3$).

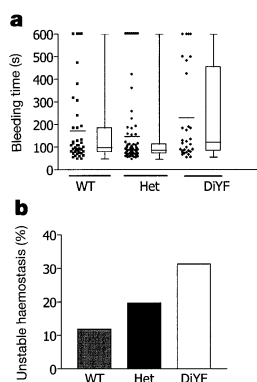


Figure 4 Bleeding times and rebleeding occurrences in diYF mutant mice. **a**, Time taken for initial cessation of bleeding. Each symbol represents a single animal (WT, squares, $n = 42$; Het, circles, $n = 72$; diYF, diamonds, $n = 32$). Box and whisker graphs indicate the median, first and third quartiles, and the total range of bleeding times for each genotype. **b**, Rebleeding occurrences in WT (grey), heterozygous (black) and diYF (white) mice. Rebleeding occurred in 31.25% of diYF mice ($n = 32$) as compared with 11.9% ($n = 42$) of WT and 19.7% ($n = 76$) of heterozygous animals.

quartile to third quartile), similar to the 100-s value reported for this model¹⁷. The rebleeding phenotype observed in the diYF mice is clearly distinct from the greatly enhanced bleeding time (Glanzmann's phenotype) observed in $\beta 3$ null mice¹⁷.

The ICY mutation in the diYF mice precludes $\beta 3$ tyrosine phosphorylation and impacts on outside-in $\alpha \text{IIb}\beta 3$ signalling, probably by depriving this integrin of its ability to interact, in a phosphorylation-dependent manner, with other proteins. Our data show the importance of $\beta 3$ outside-in signalling in key platelet functions, including formation of stable aggregates and retraction of clots, and implicate $\beta 3$ outside-in signalling in platelet pathology such as the formation of arterial thrombi that mediate myocardial infarction and stroke¹⁸. In addition, other residues within the ICY motif can impact on platelet function. In particular, mutation of Ser 752 to proline results in a Glanzmann's phenotype, disrupting both inside-out and outside-in $\alpha \text{IIb}\beta 3$ signalling^{19,20}. Our data have obvious implications for other integrins, as the ICY motif is conserved in several β -integrin subunits and across species²¹. The disruption of integrin function is an effective and widely explored therapeutic strategy and although, to date, all integrin antagonists function by blocking the binding of adhesive proteins, disruption of the ICY motif may provide attractive new mechanisms for therapeutic modification of integrin function. □

Methods

Vector construction

A 7-kilobase (kb) region of $\beta 3$ containing exons M and N (using the terminology for the human gene²²) which encode the cytoplasmic domain was used to construct the targeting vector. A *PstI/SacI* fragment containing Exon N was subcloned into pBSKSII and the two tyrosine residues were mutated to phenylalanines by site-directed mutagenesis with new restriction sites also introduced (*DraI* at Y747F, *TaqI* at Y759F). A neomycin resistance gene flanked by FRT recognition sites (from the vector pBS-FRT-neo obtained from K. Jones) was inserted into the *HindIII* site in the intron between exons M and N. This vector was transfected into ES cells, and correctly targeted clones were identified by Southern blotting and PCR. Two selected ES cell clones were then transiently transfected with the FLP recombinase²³ to remove the 1.9-kb neo cassette, leaving behind the FRT recognition site remnants with an *XbaI* restriction site (Fig. 1a, X*). Removal of neo but retention of the mutant $\beta 3$ allele was again confirmed by Southern blotting and PCR.

Preparation of washed platelets and platelet-rich plasma (PRP)

Blood obtained by cardiac puncture (700 μl), was drawn into a mix of 140 μl of ACD (2.5% trisodium citrate, 2% dextrose, 1.5% citric acid (monohydrate)), 560 μl of saline and PGE₁ (Sigma) at 50 ng ml^{-1} final concentration, and washed platelets were obtained as

described²⁴. Platelets pooled from a number of animals were resuspended in tyrodes buffer with 1 mM MgCl_2 at $2 \times 10^8 \text{ ml}^{-1}$. PRP was obtained as described²⁴, pooled and platelet number normalized for each genotype (from 1.2 to 1.8×10^8 platelets per ml in separate experiments) using platelet-poor plasma. MgCl_2 (1 mM) was added to PRP before aggregation.

$\beta 3$ expression levels

Samples corresponding to 1.5×10^7 washed platelets were lysed in a 1% TX-100 lysis buffer and analysed by western blotting using both a polyclonal serum raised against full length $\beta 3$ (#2656), and a monoclonal antibody raised against the cytoplasmic domain of $\beta 3$ (C3a.19.5). For surface expression of $\beta 3$, platelets were stained with a fluorescein isothiocyanate (FITC)- $\beta 3$ monoclonal antibody (Pharmingen) or a polyclonal anti- $\alpha \text{IIb}\beta 3$ (#41)²⁴ followed by a secondary FITC-anti-rabbit IgG (Jackson Immunoresearch) and analysed by FACSsort (Becton Dickinson).

Aggregation assays

Aggregation was measured in a Chrono-log lumiaggregometer. Platelets were induced to aggregate by addition of thrombin (washed platelets) or ADP (PRP) with stirring at 37 °C.

FITC-fibrinogen binding

PRP (10 μl) was incubated with a mixture of agonist, FITC-fibrinogen and FACS buffer (1 mM Mg -tyrodes, 0.1% BSA, 0.1% dextrose) in a final volume of 50 μl . After 10 min at room temperature, samples were fixed by the addition of paraformaldehyde (1% final volume). FACS buffer (200 μl) was then added and samples were read on a FACSsort (Becton Dickinson). In samples stimulated with thrombin, a Gly-Pro-Arg-Pro peptide (Bachem) at 2.5 mM final concentration was also present to prevent fibrin-mediated coagulation of the PRP. Fibrinogen was labelled as described²⁴ and saturating amounts, determined empirically, were added to each sample.

Clot retraction

CaCl_2 (2 mM) was added to PRP along with 1, 3 or 10 units ml^{-1} thrombin; 1 ml of this was dispensed into a glass tube containing a paper clip and samples were incubated at 37 °C. Within 20 min a fibrin clot had formed which retracted over time. After 2 h, residual plasma volumes, which indicate the degree to which retraction has occurred and which were determined by removing the clots and measuring the remaining plasma volume, were obtained where possible. However, the fragile nature of the clots at times precluded such measurements, particularly with the larger clots observed in the PRP treated with 1 unit ml^{-1} thrombin.

Tail-bleeding measurements

Eight-week-old mice, the progeny of heterozygote crosses, were anaesthetized 6 min before having 0.2 mm of the tip of their tail cut with a sharp scalpel blade. The tail was placed in a solution of saline at 37 °C and the time taken for the flow of blood to cease was recorded. Where necessary, bleeding was manually stopped at the 10-min time point to prevent death. Genotyping of the animals took place after this procedure so that tails were intact for bleeding-time measurements. If bleeding restarted within 1 min, this result was recorded as a rebleed and taken to indicate an unstable haemostatic event.

Received 28 June; accepted 27 August 1999.

- Hynes, R. O. Integrins: versatility, modulation, and signaling in cell adhesion. *Cell* **69**, 11–25 (1992).
- Shattil, S., Kashiwagi, H. & Pampori, N. Integrin signaling: the platelet paradigm. *Blood* **91**, 1–14 (1998).
- Phillips, D. R., Charo, I. E. & Scarborough, R. M. GPIIb-IIIa: the responsive integrin. *Cell* **65**, 359–362 (1991).
- Yamaguchi, K. *et al.* Flow cytometric analysis of changes in cytoskeletal proteins during platelet destruction and activation using a monoclonal antibody against platelet myosin. *Am. J. Hematol.* **44**, 106–111 (1993).
- Clark, E. A., Shattil, S. J., Ginsberg, M. H., Bolen, J. & Brugge, J. S. Regulation of the protein tyrosine kinase pp72syk by platelet agonists and the integrin $\alpha \text{IIb}\beta 3$. *J. Biol. Chem.* **269**, 28859–28864 (1994).
- Lipfert, L. *et al.* Integrin-dependent phosphorylation and activation of the protein tyrosine kinase pp125FAK in platelets. *J. Cell. Biol.* **119**, 905–912 (1992).
- Law, D. A., Nannizzi-Alaimo, L. & Phillips, D. R. Outside-in integrin signal transduction. $\alpha \text{IIb}\beta 3$ -(GP IIb IIIa) tyrosine phosphorylation induced by platelet aggregation. *J. Biol. Chem.* **271**, 10811–10815 (1996).
- Phillips, D. R., Jennings, L. K. & Edwards, H. H. Identification of membrane proteins mediating the interaction of human platelets. *J. Cell. Biol.* **86**, 77–86 (1980).
- van der Geer, P. *et al.* A conserved amino-terminal Shc domain binds to phosphotyrosine motifs in activated receptors and phosphopeptides. *Curr. Biol.* **5**, 404–412 (1995).
- Jenkins, A. *et al.* Tyrosine phosphorylation of the $\beta 3$ cytoplasmic domain mediates integrin-cytoskeletal interactions. *J. Biol. Chem.* **273**, 13878–13885 (1998).
- Guinto, E. R. *et al.* Unexpected crucial role of residue 225 in serine proteases. *Proc. Natl Acad. Sci. USA* **96**, 1852–1857 (1999).
- Filardo, E. J., Brooks, P. C., Deming, S. L., Damsky, C. & Chersesh, D. A. Requirement of the NPXY motif in the integrin $\beta 3$ subunit cytoplasmic tail for melanoma cell migration *in vitro* and *in vivo*. *J. Cell. Biol.* **30**, 441–450 (1995).
- Romzek, N. C. *et al.* Use of a $\beta 1$ integrin-deficient human T cell to identify $\beta 1$ integrin cytoplasmic domain sequences critical for integrin function. *Mol. Biol. Cell* **9**, 2715–2777 (1998).
- Ylanne, J. *et al.* Distinct functions of integrin α and β subunit cytoplasmic domains in cell spreading and formation of focal adhesions. *J. Cell. Biol.* **122**, 223–233 (1993).

15. Blystone, S. D., Williams, M. P., Slater, S. E. & Brown, E. J. Requirement of integrin $\beta 3$ tyrosine 747 for $\beta 3$ tyrosine phosphorylation and regulation of $\alpha v \beta 3$ avidity. *J. Biol. Chem.* **272**, 28757–28761 (1997).
16. Fox, J. E. *et al.* On the role of the platelet membrane skeleton in mediating signal transduction. Association of GP IIb–IIIa, pp60c-src, pp62c-yes, and the p21ras GTPase-activating protein with the membrane skeleton. *J. Biol. Chem.* **268**, 25973–25984 (1993).
17. Hodivala-Dilke, K. M. *et al.* $\beta 3$ -integrin-deficient mice are a model for Glanzmann thrombasthenia showing placental defects and reduced survival. *J. Clin. Invest.* **103**, 229–238 (1999).
18. Collier, B. S. Platelets and thrombolytic therapy. *N. Engl. J. Med.* **322**, 33–42 (1990).
19. Chen, Y. P. *et al.* Ser-752→Pro mutation in the cytoplasmic domain of integrin $\beta 3$ subunit and defective activation of platelet integrin $\alpha IIb \beta 3$ (glycoprotein IIb–IIIa) in a variant of Glanzmann thrombasthenia. *Proc. Natl Acad. Sci. USA* **89**, 10169–10173 (1992).
20. Chen, Y. P., O'Toole, T. E., Ylanne, J., Rosa, J. P. & Ginsberg, M. H. A point mutation in the integrin $\beta 3$ cytoplasmic domain (S752→P) impairs bidirectional signaling through $\alpha IIb \beta 3$ (platelet glycoprotein IIb–IIIa). *Blood* **84**, 1857–1865 (1994).
21. Williams, M. J., Hughes, P. E., O'Toole, T. E. & Ginsberg, M. H. The inner world of cell adhesion: integrin cytoplasmic domains. *Trends Cell Biol.* **4**, 109–112 (1994).
22. Fitzgerald, L. A., Steiner, B., Rall, S. C. Jr, Lo, S. S. & Phillips, D. R. Protein sequence of endothelial glycoprotein IIIa derived from a cDNA clone. Identity with platelet glycoprotein IIIa and similarity to "integrin". *J. Biol. Chem.* **262**, 3936–3939 (1987).
23. Dymecki, S. M. Flp recombinase promotes site-specific DNA recombination in embryonic stem cells and transgenic mice. *Proc. Natl Acad. Sci. USA* **93**, 6191–6196 (1996).
24. Law, D. A. *et al.* Genetic and pharmacologic analyses of Syk function in $\alpha IIb \beta 3$ signaling in platelets. *Blood* **93**, 2645–2652 (1999).

Acknowledgements

We thank R. Wong for his help with the figures.

Correspondence and requests for material should be addressed to D.R.P. (e-mail: david_phillips@corr.com).

The Toll-like receptor 2 is recruited to macrophage phagosomes and discriminates between pathogens

David M. Underhill*, Adrian Ozinsky*, Adeline M. Hajjar, Anne Stevens, Christopher B. Wilson, Michael Bassetti & Alan Aderem

Department of Immunology, University of Washington, H-574 Health Sciences, Box 357650, Seattle, Washington 98195, USA

* These authors contributed equally to this work

Macrophages orchestrate innate immunity by phagocytosing pathogens and coordinating inflammatory responses¹. Effective defence requires the host to discriminate between different pathogens. The specificity of innate immune recognition in *Drosophila* is mediated by the Toll family of receptors^{2,3}; Toll mediates anti-fungal responses, whereas 18-wheeler mediates anti-bacterial defence^{4–6}. A large number of Toll homologues have been identified in mammals, and Toll-like receptor 4 is critical in responses to Gram-negative bacteria^{7–11}. Here we show that Toll-like receptor 2 is recruited specifically to macrophage phagosomes containing yeast, and that a point mutation in the receptor abrogates inflammatory responses to yeast and Gram-positive bacteria, but not to Gram-negative bacteria. Thus, during the phagocytosis of pathogens, two classes of innate immune receptors cooperate to mediate host defence: phagocytic receptors, such as the mannose receptor, signal particle internalization, and the Toll-like receptors sample the contents of the vacuole and trigger an inflammatory response appropriate to defence against the specific organism.

Macrophage phagocytosis of the yeast cell-wall particle zymosan is accompanied by secretion of the inflammatory mediator tumour necrosis factor- α (TNF- α). Internalization of zymosan is mediated, in part, by the mannose receptor, a classic example of an innate immune pattern-recognition molecule, but the signalling pathway

that elicits TNF- α production is unknown¹². Given the role of Toll-like receptors in pro-inflammatory signalling during innate defence, we explored whether these molecules are also found on phagosomes. We expressed haemagglutinin A (HA)-tagged murine Toll-like receptor 2 (TLR2) in the mouse macrophage cell line, RAW-TT10. In the absence of a phagocytic stimulus, the receptor was distributed uniformly on the cell surface (Fig. 1a). After a 5-min incubation with zymosan particles, TLR2 was highly enriched on phagosomes (Fig. 1b, c). Phagocytosis occurs through a series of discrete steps: initially the particle is bound to the cell surface; actin polymerization leads to pseudopod extension around the particle; the particle is completely engulfed by the cell; and lastly actin is depolymerized, allowing further phagosome maturation through membrane-transport events¹. The enrichment of TLR2 was apparent throughout all phases of internalization (Fig. 1b, c, zymosan particles 1–4): the receptor was enriched under bound particles (zymosan 1), in processes extending around extracellular particles (zymosan 2) and in membranes around fully engulfed particles (zymosan 3), and it remained enriched on the phagosome for several minutes after actin disassembly (zymosan 4; and data not shown). Within 10 min after internalization, however, TLR2 was no longer enriched on the vacuole (data not shown).

Phagocytosis of particles is tightly coupled to the production of pro-inflammatory cytokines, but the mechanism by which this

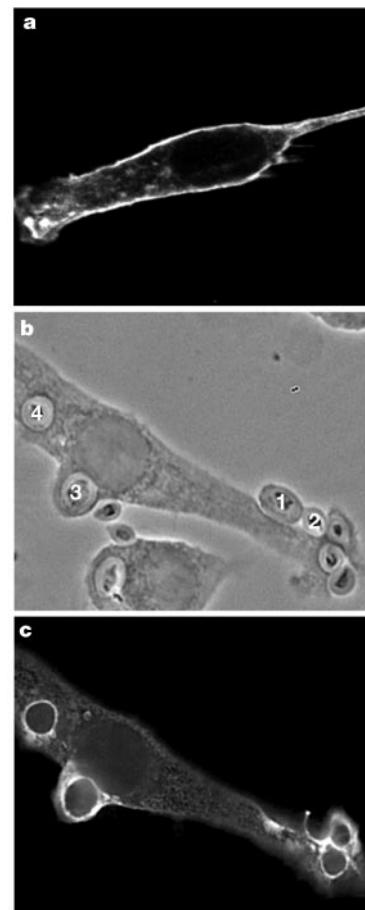


Figure 1 TLR2 is enriched on phagosomes containing yeast particles. **a**, Epitope-tagged TLR2 was detected on the cell surface by immunofluorescence microscopy in RAW-TT10 cells transiently expressing the receptor. **b**, When these cells were incubated with zymosan for 5 min and processed for immunofluorescence microscopy, zymosan particles, both external (particles 1 and 2) and internal (particles 3 and 4), were clearly visible by phase microscopy. **c**, TLR2 was markedly enriched around the particles during all stages of internalization.

occurs is unknown. The presence of TLR2 on phagosomes indicated that it might mediate pro-inflammatory signals during zymosan phagocytosis. To establish a role for TLR2 in zymosan-induced cytokine release, we needed to define a dominant-negative form of this receptor. Genetic evidence in the C3H/HeJ mouse indicates that a P712H substitution in the related Toll-like receptor 4 (TLR4) may act as a dominant-negative mutation of the receptor and prevent LPS signalling in the mouse^{10,11,13}. We therefore constructed a putative dominant-negative mutant TLR2, in which proline 681 was replaced by a histidine (TLR2-P681H), a mutation that is equivalent to the P712H substitution in TLR4. Chinese hamsters fail to express functional TLR2 because of a frame-shift mutation in the receptor¹⁴, making Chinese hamster ovary (CHO) cells an ideal model system in which to define dominant-negative mutants of TLR2 signalling. Expression of TLR2 in CHO cells resulted in a moderate level of constitutive activation of NF- κ B, which was not increased substantially upon exposure of the cells to zymosan (Fig. 2a). CD14 has been shown to facilitate the activation of TLR2 by bacterial products^{15,16}. Cotransfection of CHO cells with CD14 and TLR2 greatly enhanced zymosan-induced NF- κ B activation (Fig. 2a), whereas expression of CD14 by itself did not mediate zymosan-induced signalling (data not shown). TLR2-P681H did not mediate zymosan-induced activation of NF- κ B in the presence of CD14, indicating that the mutation abrogated the capacity of the receptor to signal (Fig. 2a). Importantly, TLR2-P681H acted as a dominant-negative inhibitor of TLR2-mediated activation of NF- κ B (Fig. 2a). Identical results were obtained with *Staphylococcus aureus* (data not shown). When expressed in RAW-TT10 macrophages, TLR2-P681H was distributed at the cell surface and was substantially enriched around zymosan-containing phagosomes, indicating that it was properly localized to function as a dominant-

negative inhibitor of particle-induced signalling through endogenous TLR2 (Fig. 2b).

To determine whether TLR2-P681H was capable of inhibiting the production of pro-inflammatory cytokines during phagocytosis in macrophages, we designed a single-cell assay in which we could simultaneously analyse the expression of the mutant receptor and zymosan-induced TNF- α production. In this assay, transient expression of TLR2-P681H and green fluorescent protein (GFP) from a single bicistronic transcript permitted GFP intensity (as measured by flow cytometry) to reflect accurately the level of TLR2-P681H expression (see Supplementary Information). RAW-TT10 cells were transfected with a control vector expressing only GFP, and 100,000 cells were analysed by flow cytometry. TNF- α production remained at background levels over three logs of expression of GFP (Fig. 2c). When stimulated with zymosan particles, TNF- α production was induced equally in all cells irrespective of their level of GFP expression (Fig. 2d). Expression of TLR2-P681H correlated in a dose-dependent manner with the inhibition of TNF- α production induced by zymosan (Fig. 2e). Cells producing TLR2-P681H (as measured by GFP expression) clearly failed to produce TNF- α , whereas cells in the same sample that did not express the inhibitory receptor produced TNF- α . Expression of TLR2-P681H in RAW-TT10 cells also inhibited zymosan-induced activation of NF- κ B, as measured by its translocation from the cytosol to the nucleus (see Supplementary Information). Particle binding and internalization were unaffected by the expression of the mutant receptor (data not shown; and see below), indicating that phagocytosis and particle-induced cytokine production were uncoupled in these cells. The full spectrum of TNF- α production by GFP-positive or GFP-negative cells (Fig. 2e, boxed regions) is shown in Fig. 2f. GFP-negative cells (Fig. 2f, thin line) had appreciable zymosan-induced TNF- α

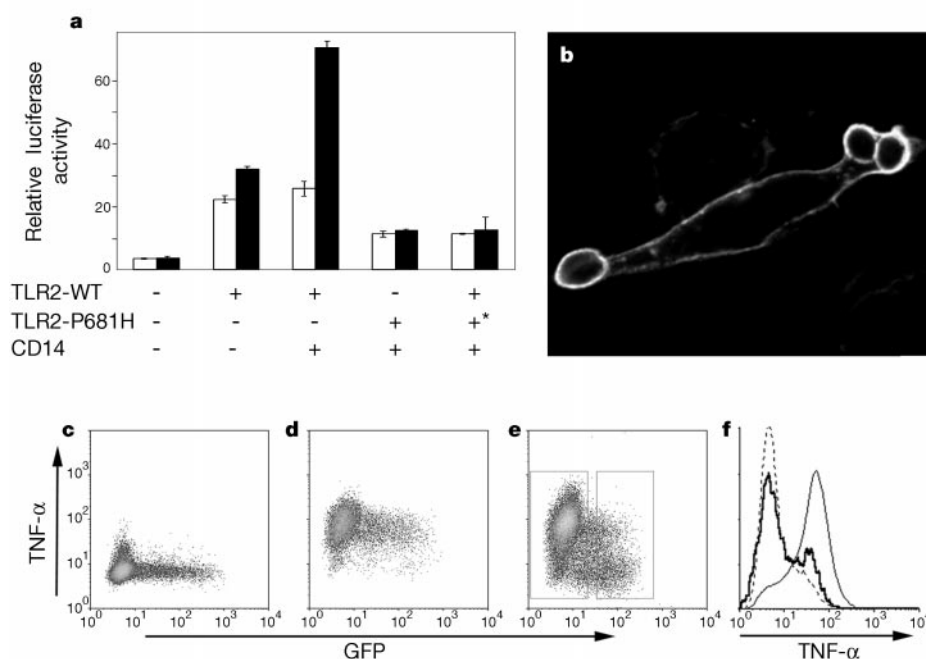


Figure 2 Expression of TLR2-P681H blocks zymosan-induced TNF- α production. **a**, The relative activity of a NF- κ B-luciferase reporter was assessed in CHO cells transiently expressing the indicated constructs before (open bars) or after (filled bars) stimulation with zymosan. In the sample showing inhibition of TLR2 by TLR2-P681H (asterisk), the mutant receptor was cotransfected with the wild-type receptor at a 10:1 ratio. **b**, Epitope-tagged TLR2-P681H, transiently expressed in RAW-TT10 cells, was highly enriched on zymosan phagosomes. The HA tag was detected by immunofluorescence microscopy as in Fig. 1. **c**, **d**, RAW-TT10 cells transiently transfected with the control expression vector (pTIGZ2+)

were stained for production of TNF- α and analysed by flow cytometry following no stimulation (**c**), or after incubation with zymosan (**d**). **e**, Cells expressing TLR2-P681H failed to induce TNF- α after stimulation with zymosan. **f**, Using the regions defined in **e**, the thin line indicates TNF- α produced in stimulated GFP-low cells, and the bold line indicates TNF- α produced in stimulated, TLR2-P681H-expressing cells (GFP-high). The dotted line indicates background, unstimulated TNF- α production; y axis, relative cell number.

production, which was inhibited in TLR2-P681H-expressing cells (thick line), effectively to unstimulated levels (dotted line). Thus, TLR2 couples the phagocytosis of zymosan to TNF- α production, and the extent of the inhibition (>90%) suggests that this is the primary pathway.

As phagocytosis of yeast resulted in TLR2-mediated signalling, we examined the potential for this receptor to signal TNF- α production during the phagocytosis of other pathogens. TNF- α production induced during the phagocytosis of *S. aureus*, a Gram-positive bacterium, was completely inhibited by the expression of TLR2-P681H (Fig. 3), as was TNF- α induction by the Gram-positive bacterial cell-wall components lipoteichoic acid and peptidoglycan (data not shown). By contrast, TNF- α production elicited during the phagocytosis of *Salmonella minnesota* R595, a Gram-negative bacterium, was unaffected by the expression of TLR2-P681H (Fig. 3). Similarly, lipopolysaccharide (LPS), a key pro-inflammatory component of the Gram-negative bacterial cell wall, induced TNF- α normally in the presence of TLR2-P681H (Fig. 3). Thus, TLR2 discriminates yeast and Gram-positive bacteria from Gram-negative bacteria. The C3H/HeJ mouse, which bears a P712H mutation in TLR4, is hyporesponsive to LPS, indicating that TLR4 participates in the recognition of Gram-negative bacteria^{10,11}. Expression of TLR4-P712H in RAW-TT10 macrophages strongly inhibited LPS-induced TNF- α production, whereas it had no effect on *S. aureus*-induced signalling (Fig. 3). TLR4-P712H inhibited *S. minnesota*-induced TNF- α production to a lesser degree than was observed with soluble LPS (Fig. 3). These data show that individual Toll family receptors convey specificity in mediating pro-inflammatory signals from different pathogens.

Further insight into the mechanism by which TLR2 signals was obtained by the observation that expression of the wild-type

receptor partially inhibited zymosan- and *S. aureus*-induced TNF- α production, but had no effect on *S. minnesota*- and LPS-induced TNF- α production (Fig. 3; and data not shown). The data indicate that over-expression of the wild-type receptor may dilute out other molecules required for TLR2 function, perhaps a structural component of the receptor, and further shows that TLR2 does not participate in the response of macrophages to Gram-negative bacteria.

MyD88 is a cytoplasmic adaptor protein that binds to the Toll-homology domain of TLR4 and the interleukin-1 receptor (IL-1R), and mediates their downstream signalling^{17–20}. Binding of MyD88 to these receptors is mediated by a C-terminal Toll-homology domain, whereas its N-terminal death domain is important for the propagation of the signal through the protein kinase IRAK (IL-1 receptor associated kinase). The C-terminal Toll domain of MyD88 is a dominant-negative inhibitor of TLR4 and IL-1R signalling, whereas the death domain constitutively activates these pathways^{17–19}. Expression of the C-terminal MyD88 dominant negative (MyD88-DN) correlated precisely with inhibition of TNF- α production induced during the phagocytosis of zymosan and the Gram-positive bacterium, *S. aureus* (Fig. 4a). The failure of TNF- α induction could not be attributed to an inhibition of particle binding or internalization, as the internalization of fluorescently labelled zymosan, *S. aureus* and *S. minnesota* was unaffected by MyD88-DN expression (Fig. 4b). Thus, MyD88 is required for TLR2 signalling in the macrophage. MyD88-DN also inhibited TNF- α production induced by the Gram-negative bacterium *S. minnesota* and by LPS (Fig. 4a), although the inhibitor was less effective against the whole bacterium (as observed with TLR4-P712H in Fig. 3). These later data indicate that, during phagocytosis of Gram-negative bacteria, a MyD88-dependent pathway (presumably initiated by TLR4) and a MyD88-independent pathway may contribute to the complete induction of TNF- α . This hypothesis concurs with the observation that, although the TLR4 mutation in the C3H/HeJ mouse blocks soluble LPS signalling, Gram-negative bacteria still stimulate cytokine production (perhaps through Toll-independent pathways)²¹. Thus, both genetic data and our results indicate that TLR4 may be the primary mediator of LPS signalling, whereas our data show that TLR2 principally mediates signals from yeast and Gram-positive bacteria in macrophages. Previously,

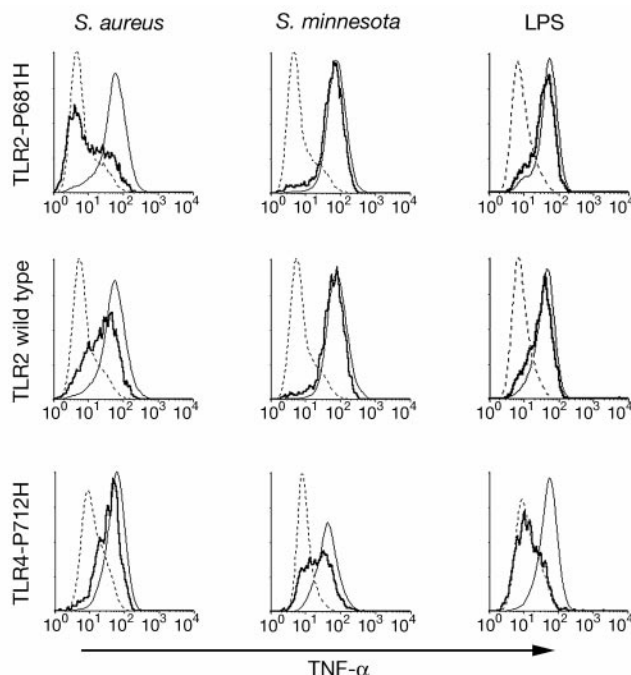


Figure 3 Inhibition of particle-induced TNF- α production by TLR2-P681H is selective. RAW-TT10 cells transiently transfected with TLR2-P681H (upper panels), wild-type TLR2 (middle panels) or TLR4-P712H (lower panels) were analysed for particle-induced TNF- α production. The cells were stimulated as indicated with *S. aureus*, *S. minnesota* or LPS. Using the regions defined in Fig. 2e, the thin lines indicate TNF- α produced in stimulated, GFP-low cells, and the bold lines indicate TNF- α produced in stimulated, TLR-expressing cells (GFP-high). The dotted lines indicate background, unstimulated TNF- α production; y axes, relative cell number.

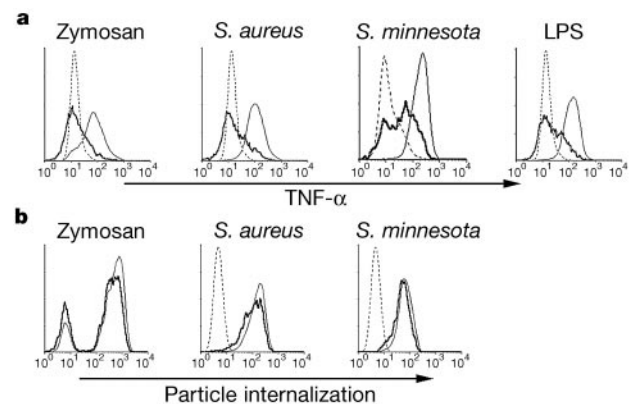


Figure 4 Expression of MyD88-DN blocks particle-induced TNF- α production but not particle internalization. **a**, RAW-TT10 cells transiently expressing MyD88-DN were stimulated as indicated with zymosan, *S. aureus*, *S. minnesota* or LPS, and TNF- α production was measured. Using the regions as defined in Fig. 2e, the thin lines indicate TNF- α produced in stimulated, GFP-low cells, and the bold lines indicate TNF- α produced in stimulated, MyD88-DN-expressing cells (GFP-high). The dotted lines indicate background, unstimulated TNF- α production. **b**, Cells transiently expressing MyD88-DN were incubated with fluorescently labelled zymosan, *S. aureus*, or *S. minnesota* as indicated, and particle internalization was assessed by flow cytometry; y axes, relative cell number.

human TLR2 has been implicated in LPS signalling in non-macrophage cell lines, and it not clear why our data, obtained in cells that are naturally LPS-responsive, differ^{22,23}; however, our data concur with genetic evidence in mice and hamsters, in which a mutation in TLR4 abrogates LPS signalling whereas a mutation in TLR2 does not^{10,11,14}. We have also shown that the natural mutation in the TLR4 gene in the C3H/HeJ LPS-hyporesponsive mouse (P712H) acts as a dominant-negative inhibitor of TLR4 function. Notably, the corresponding proline-to-histidine mutation in TLR2 (P681H) also acts as a dominant-negative mutation in TLR2. This proline is conserved in all known mammalian Toll receptors, except TLR3 (ref. 7–9), and might couple Toll receptors to a common signalling pathway.

Macrophages are capable of detecting the nature of the pathogen within the phagocytic vacuole, but the mechanisms by which this occurs are unknown¹. Our data indicate that the toll-like receptors may be recruited to phagosomes, where they sample the contents and determine the nature of the pathogen; TLR2 detects Gram-positive bacteria and fungi, whereas TLR4 detects Gram-negative bacteria. Thus, specific Toll-like receptors may distinguish the contents of the phagosome and participate in the elaboration of an inflammatory response appropriate for defence against a specific pathogen. □

Methods

Murine TLR2 cloning

Two mouse ESTs (Genbank AA863729 and D77677) with homology to the human TLR2 sequence were identified, and a 0.9-kilobase (kb) complementary DNA was generated by reverse transcription of RAW 264.7 total RNA using a primer spanning the stop codon in the first EST (5'-GGTGGAGAACCTAGGACTTTATTGC-3'). Polymerase chain reaction (PCR) was carried out with the same primer, and the 5' most homologous sequences in the second EST (5'-GCTCTATTTCCAGAAATAAGCTG-3'). To obtain the complete open reading frame (ORF), 5' rapid amplification of complementary DNA ends (RACE) was performed on RAW 264.7 RNA (Gibco-BRL). Products from 0.5 to 1.5 kb were amplified, cloned and sequenced, completing the murine TLR2 ORF (Genbank AF185284).

DNA expression vectors

The mammalian expression vector, pNeo/Tak, was constructed to combine expression of the tetracycline transactivator²⁴ from a tetracycline-regulated promoter (pTet-Tak, Gibco-BRL) with neomycin selection. The neomycin resistance marker was from pcDNA3 (Invitrogen), and the remainder of the plasmid backbone was derived from pBluescript SK (Stratagene).

The expression vector pTIGZ2+ was assembled to direct tetracycline-regulated expression of a protein from the same messenger RNA transcript as GFP. In this vector, the tetracycline-regulated promoter from pTetSplice (Gibco-BRL) directed expression of a mRNA encoding the protein of interest followed by the cap-independent translational enhancer region of pCITE (Novagen) driving translation of enhanced GFP (from pGFP-N1, Clontech). The polyadenylation signal and remainder of the vector backbone were from pcDNA3.1/Zeo (Invitrogen). During construction, the tetracycline-inducible promoter was modified to contain two binding sites for the tetracycline transactivator instead of seven.

The HA-TLR2 expression vector was constructed by amplifying the coding region of mouse TLR2 from amino acids R21–S784 (removing the endogenous signal peptide) by PCR, and cloning the fragment into a modified pDisplay (Invitrogen) which provides a signal peptide and a haemagglutinin A epitope. pDisplay had been modified by deletion of the PDGF transmembrane domain and replacement of the neomycin resistance cassette by one with zeocin resistance (pcDNA3.1/Zeo, Invitrogen). For single-cell TNF- α measurements, the HA-tagged TLR2 coding region was transferred from pDisplay to TIGZ2+. The TLR2 dominant-negative P681H mutation was generated by PCR and confirmed by sequencing. The MyD88 dominant-negative construct was generated by amplifying the MyD88 C-terminal coding region (amino acids 146–296) by PCR from cDNA of RAW 264.7 cells and cloning the fragment into TIGZ2+. The murine TLR4 coding region was amplified by RT-PCR from RAW cells (Genbank AF185285), and cloned into pTIGZ2+. The dominant-negative P712H mutation was generated by PCR and confirmed by sequencing, and the resulting coding region was cloned into pTIGZ2+. Murine CD14 was amplified by RT-PCR from RAW cells and cloned into the expression vector pcDNA3.1/Zeo (Invitrogen).

Transfection

RAW 264.7 (ATCC# TIB-71) and CHO-K1 (ATCC# CRL-9618) cells were transfected as described²⁵. All experiments were performed 24 h after transfection. For generation of a tetracycline transactivator-expressing RAW cell line, RAW 264.7 cells were transfected with pNeo/Tak, and stable cell lines were selected using 400 $\mu\text{g ml}^{-1}$ G418 (Gibco-BRL). After 10 days of selection, the cells were cloned by limiting dilution, and one cell line (designated

RAW-TT10) that showed good tetracycline-regulated expression from a subsequently transfected reporter plasmid was used for all experiments. Tetracycline was always absent from the media, resulting in strong activity of the tetracycline-regulated promoter.

Immunofluorescence microscopy

RAW cells were processed for immunofluorescence microscopy as described²⁵. Cells grown on glass coverslips were incubated for 5 min at 37 °C with zymosan (Molecular Probes), fixed and permeabilized, and nonspecific sites were blocked by incubation in blocking buffer (PBS, 10% calf serum, 0.1% ovalbumin and 5 mM sodium azide). HA-tagged TLR2 was detected using a mouse monoclonal antibody (HA.11, Babco) precleared against zymosan. Specific binding was detected with fluorescein-conjugated goat anti-mouse secondary antibody (Jackson ImmunoResearch). After mounting, the cells were observed by fluorescence microscopy using an Axiocvert TV microscope (Carl Zeiss) equipped with a cooled CCD camera (Princeton Instruments) and Metamorph imaging software (Universal Imaging).

Luciferase assays

CHO-K1 cells were transiently transfected with 2 μg of NF- κB reporter construct (ELAM-1 firefly luciferase²⁶) and 0.2 μg of a construct directing expression of *Renilla* luciferase under control of the constitutively active β -actin promoter²⁷, together with 3 μg of each of the expression constructs indicated in the text, and plated into 24-well tissue-culture plates. Twenty-four hours after transfection, the cells were stimulated with 3×10^6 zymosan particles (Molecular Probes) per well for 5 h, and luciferase activity was measured using the Dual-Luciferase Reporter Assay System (Promega) according to the manufacturer's instructions. Data are presented as the mean $\pm$ standard deviation of triplicate samples and are representative of three independent experiments.

Detection of intracellular TNF- α

S. aureus (clinical isolate) and *S. minnesota* (ATCC#49284) were grown to saturation, heat killed and stored as frozen aliquots. For particle stimulation, the RAW cells were incubated for 30 min at 37 °C with 3×10^6 zymosan (Molecular Probes), *S. aureus* or *S. minnesota* particles per well. After exposure to the particles, the cells were incubated for 4 h at 37 °C in the presence of 5 $\mu\text{g ml}^{-1}$ brefeldin A to accumulate intracellular TNF- α . Where indicated, cells were exposed to 10 ng ml^{-1} LPS (*S. minnesota* R595, List), 10 $\mu\text{g ml}^{-1}$ lipoteichoic acid (*S. aureus*, Sigma) or 10 $\mu\text{g ml}^{-1}$ peptidoglycan (*S. aureus*, Fluka) with 5 $\mu\text{g ml}^{-1}$ brefeldin A for 4 h at 37 °C. After blocking Fc-receptors with 2.4G2 hybridoma supernatant, the cells were fixed with 4% paraformaldehyde in PBS. The cells were permeabilized and stained with phycoerythrin-conjugated anti-mouse TNF- α (Pharmingen) diluted in 1% fetal calf serum and 0.1% saponin in PBS. After two washes, the cells were analysed by flow cytometry using a FACScan (Beckton Dickinson) and WinMDI software (Joseph Trotter, Scripps). Brefeldin A alone did not stimulate TNF- α production (data not shown).

Phagocytosis assay

Phagocytosis by RAW cells was measured by flow cytometry of cells that had ingested fluorescently labelled particles²⁵. Zymosan, *S. aureus* or *S. minnesota* particles (3×10^6) labelled with tetramethylrhodamine were added to the macrophage monolayers and settled onto the cells by centrifugation. The cells were allowed to internalize the particles for 10 min at 37 °C, and uninternalized particles were washed away with PBS. For cells ingesting zymosan, remaining extracellular particles were dissolved by addition of 100 U ml^{-1} lyticase (Sigma) in PBS for 10 min at room temperature. The cells were resuspended in PBS + 1 mM EDTA, fixed in 1% formaldehyde, and analysed by flow cytometry. For cells ingesting bacteria, uninternalized bacteria were stripped from the cell surface by incubation in trypsin/EDTA for 5 min at room temperature, and the cells were fixed and analysed as above. It was confirmed by microscopy that the only remaining cell-associated fluorescently labelled particles had been internalized.

Received 19 August; accepted 10 September 1999.

1. Aderem, A. & Underhill, D. M. Mechanisms of phagocytosis in macrophages. *Annu. Rev. Immunol.* **17**, 593–623 (1999).
2. Medzhitov, R. & Janeway, C. A. Jr Self-defence: the fruit fly style. *Proc. Natl Acad. Sci. USA* **95**, 429–430 (1998).
3. Hoffmann, J. A., Kafatos, F. C., Janeway, C. A. & Ezekowitz, R. A. Phylogenetic perspectives in innate immunity. *Science* **284**, 1313–1318 (1999).
4. Lemaître, B., Reichhart, J. M. & Hoffmann, J. A. *Drosophila* host defense: differential induction of antimicrobial peptide genes after infection by various classes of microorganisms. *Proc. Natl Acad. Sci. USA* **94**, 14614–14619 (1997).
5. Eldon, E. *et al.* The *Drosophila* 18 wheeler is required for morphogenesis and has striking similarities to Toll. *Development* **120**, 885–899 (1994).
6. Williams, M. J., Rodriguez, A., Kimbrell, D. A. & Eldon, E. D. The 18-wheeler mutation reveals complex antibacterial gene regulation in *Drosophila* host defense. *EMBO J.* **16**, 6120–6130 (1997).
7. Chaudhary, P. M. *et al.* Cloning and characterization of two Toll/Interleukin-1 receptor-like genes TIL3 and TIL4: evidence for a multi-gene receptor family in humans. *Blood* **91**, 4020–4027 (1998).
8. Rock, F. L., Hardiman, G., Timans, J. C., Kastelein, R. A. & Bazan, J. F. A family of human receptors structurally related to *Drosophila* Toll. *Proc. Natl Acad. Sci. USA* **95**, 588–593 (1998).
9. Takeuchi, O. *et al.* TLR6: a novel member of an expanding Toll-like receptor family. *Gene* **231**, 59–65 (1999).
10. Poltorak, A. *et al.* Defective LPS signaling in C3H/HeJ and C57BL/10ScCr mice: mutations in Tlr4 gene. *Science* **282**, 2085–2088 (1998).
11. Qureshi, S. T. *et al.* Endotoxin-tolerant mice have mutations in toll-like receptor 4 (Tlr4). *J. Exp. Med.* **189**, 615–625 (1999).

12. Stahl, P. D. & Ezekowitz, R. A. The mannose receptor is a pattern recognition receptor involved in host defense. *Curr. Opin. Immunol.* **10**, 50–55 (1998).
13. Hoshino, K. *et al.* Cutting edge: toll-like receptor 4 (TLR4)-deficient mice are hyporesponsive to lipopolysaccharide: evidence for TLR4 as the Lps gene product. *J. Immunol.* **162**, 3749–3752 (1999).
14. Heine, H. *et al.* Cutting edge: cells that carry a null allele for toll-like receptor 2 are capable of responding to endotoxin. *J. Immunol.* **162**, 6971–6975 (1999).
15. Yoshimura, A. *et al.* Cutting edge: recognition of Gram-positive bacterial cell wall components by the innate immune system occurs via toll-like receptor 2. *J. Immunol.* **163**, 1–5 (1999).
16. Schwandner, R., Dziarski, R., Wesche, H., Rothe, M. & Kirschning, C. J. Peptidoglycan- and lipoteichoic acid-induced cell activation is mediated by toll-like receptor 2. *J. Biol. Chem.* **274**, 17406–17409 (1999).
17. Medzhitov, R. *et al.* MyD88 is an adaptor protein in the hToll/IL-1 receptor family signaling pathways. *Mol. Cell.* **2**, 253–258 (1998).
18. Muzio, M., Ni, J., Feng, P. & Dixit, V. M. IRAK (Pelle) family member IRAK-2 and MyD88 as proximal mediators of IL-1 signalling. *Science* **278**, 1612–1615 (1997).
19. Wesche, H., Henzel, W. J., Schillinglaw, W., Li, S. & Cao, Z. MyD88: an adapter that recruits IRAK to the IL-1 receptor complex. *Immunity* **7**, 837–847 (1997).
20. Adachi, O. *et al.* Targeted disruption of the MyD88 gene results in loss of IL-1- and IL-18-mediated function. *Immunity* **9**, 143–150 (1998).
21. Freudenberg, M. A. & Galanos, C. Tumor necrosis factor α mediates lethal activity of killed gram-negative and gram-positive bacteria in D-galactosamine-treated mice. *Infect. Immun.* **59**, 2110–2115 (1991).
22. Kirschning, C. J., Wesche, H., Merrill Ayres, T. & Rothe, M. Human toll-like receptor 2 confers responsiveness to bacterial lipopolysaccharide. *J. Exp. Med.* **188**, 2091–2097 (1998).
23. Yang, R. B. *et al.* Toll-like receptor-2 mediates lipopolysaccharide-induced cellular signalling. *Nature* **395**, 284–288 (1998).
24. Gossen, M. & Bujard, H. Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proc. Natl Acad. Sci. USA* **89**, 5547–5551 (1992).
25. Underhill, D. M., Chen, J., Allen, L. A. & Aderem, A. MacMARCKS is not essential for phagocytosis in macrophages. *J. Biol. Chem.* **273**, 33619–33623 (1998).
26. Schindler, U. & Baichwal, V. R. Three NF- κ B binding sites in the human E-selectin gene required for maximal tumor necrosis factor α -induced expression. *Mol. Cell. Biol.* **14**, 5820–5831 (1994).
27. Sweetser, M. T. *et al.* The roles of nuclear factor of activated T cells and ying-yang 1 in activation-induced expression of the interferon- γ promoter in T cells. *J. Biol. Chem.* **273**, 34775–34783 (1998).

Supplementary information is available in Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

Acknowledgements

Supported in part by grants from the NIH (A.A., C.B.W. and A.S.) and from the cystic fibrosis foundation (C.B.W. and A.H.). D.M.U. is an Irvington Institute postdoctoral fellow.

Correspondence and requests for materials should be addressed to A.A. (e-mail: aaderem@u.washington.edu).

Accumulation of cyclin B1 requires E2F and cyclin-A-dependent rearrangement of the anaphase-promoting complex

Claudia Lukas, Claus Storgaard Sørensen, Edgar Kramer*, Eric Santoni-Rugiu, Claes Lindenberg, Jan-Michael Peters*, Jiri Bartek & Jiri Lukas

Institute of Cancer Biology, Danish Cancer Society, Strandboulevarden 49, DK-2100, Copenhagen Ø, Denmark

* Research Institute of Molecular Pathology, Dr. Bohr-Gasse 7, A-1030 Vienna, Austria

In mammalian somatic-cell cycles, progression through the G1-phase restriction point and initiation of DNA replication are controlled by the ability of the retinoblastoma tumour-suppressor protein (pRb) family to regulate the E2F/DP transcription factors^{1,2}. Continuing transcription of E2F target genes beyond the G1/S transition is required for coordinating S-phase progression with cell division^{3–5}, a process driven by cyclin-B-dependent kinase^{6,7} and anaphase-promoting complex (APC)-mediated

proteolysis⁸. How E2F-dependent events at G1/S transition are orchestrated with cyclin B and APC activity remains unknown. Here, using an *in vivo* assay to measure protein stability in real time during the cell cycle, we show that repression of E2F activity or inhibition of cyclin-A-dependent kinase in S phase triggers the destruction of cyclin B1 through the re-assembly of APC, the ubiquitin ligase that is essential for mitotic cyclin proteolysis⁹, with its activatory subunit Cdh1 (refs 10–13). Phosphorylation-deficient mutant Cdh1 or immunodepletion of cyclin A resulted in assembly of active Cdh1–APC even in S-phase cells. These results implicate an E2F-dependent, cyclin A/Cdk2-mediated phosphorylation of Cdh1 in the timely accumulation of cyclin B1 and the coordination of cell-cycle progression during the post-restriction point period.

To elucidate the link between the pRb/E2F pathway and the key factors required for G2/M transition, we examined cyclin B1 in human U-2-OS cells (clone A5C1) conditionally expressing pRb Δ cdk, a constitutively active allele of pRb that represses transcription of E2F target genes^{5,14}. In cells synchronized in S phase by the DNA polymerase- α inhibitor aphidicolin, cyclin B1 protein was stable and accumulated to high levels which could not be further increased by addition of the proteasome-inhibiting drug, LLnL (Fig. 1a, lanes 1–3). Cdc2 was largely held in its inactive, slow-migrating form, indicating inhibitory phosphorylation on Thr 14 and Tyr 15, a modification that requires association with cyclins^{6,15} (Fig. 1a, lanes 2–3). Induction of pRb Δ cdk resulted in a fivefold reduction of cyclin B1 protein abundance, accompanied by loss of the inhibitory phosphorylation of Cdc2, and addition of LLnL reversed this effect (Fig. 1a, lanes 4–5).

Immunofluorescence analysis *in situ* confirmed that the abundant cytoplasmic pool of cyclin B1 in S-phase cells disappeared

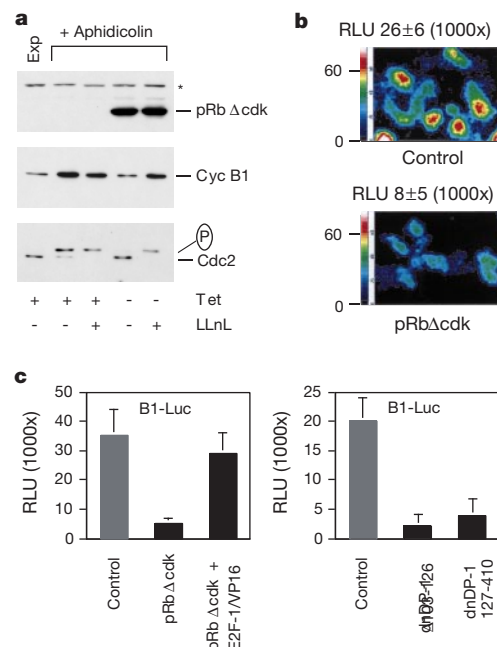


Figure 1 Repression of E2F during S phase induces proteasome-dependent degradation of cyclin B1. **a**, Immunoblots of pRb Δ cdk (induced by tetracycline removal, Tet), and endogenous cyclin B1 and Cdc2 in exponentially growing (Exp), and aphidicolin-arrested A5C1 cells, cultured with or without the proteasome inhibitor LLnL. Asterisk, cross-reacting band. **b**, Stability of B1-Luc 18 h after micro-injection of the reporter construct into aphidicolin-synchronized U-2-OS cells, measured as luciferase-mediated light emission from injected live cells; RLU, relative light units. Scale bar, 10 μ m. **c**, Summary of the *in vivo* measurements described in **b** after co-injection with B1-Luc and the indicated expression plasmids. Aphidicolin was included in all experiments.

upon induction of pRb Δ cdk, and was restored upon addition of LLN1 (data not shown). These experiments indicated that, during S-phase, pRb must remain in its phosphorylated, inactive form, allowing transactivation of E2F targets to maintain the stability of the mitotic cyclins.

To study this dynamic process in living cells, we designed a noninvasive assay for monitoring cyclin B1 protein stability quantitatively in real time. Because chimaeric proteins containing the cyclin B destruction box mimic endogenous cyclin oscillations^{16,17}, we fused the N terminus of cyclin B1 to the firefly luciferase gene, expressed this chimera from a heterologous SV40 promoter, and studied its abundance as a function of photon emission catalysed by luciferase. A combined micro-injection/*in vivo* reporter assay¹⁴ showed that co-expression of pRb Δ cdk in S-phase cells considerably reduced the light emission from B1-Luc (Fig. 1b), but not from untagged luciferase (data not shown). This indicated that, analogous to endogenous cyclin B1, protein turnover of B1-Luc may be accelerated in the presence of activated pRb, a conclusion verified by pulse-chase experiments in S-phase cells (data not shown). Two experiments supported our hypothesis that the pRb Δ cdk-induced destruction of cyclin B1 is mediated by repression of cellular E2F activity. First, co-expression of a chimeric E2F-1/VP-16 transcription factor which retains the DNA-binding specificity of E2F but cannot be sequestered by pRb¹⁸ restored the accumulation of the B1-Luc protein in S-phase cells exposed to pRb Δ cdk (Fig. 1c, left panel). Second, expression of either of two distinct dominant-negative alleles of DP-1 (dnDP-1)^{14,19}, the heterodimerizing partner of E2F, conferred the same degree of cyclin B1 instability as pRb Δ cdk (Fig. 1c, right panel), further implicating an essential E2F-dependent step in the regulation of cyclin B1 protein turnover.

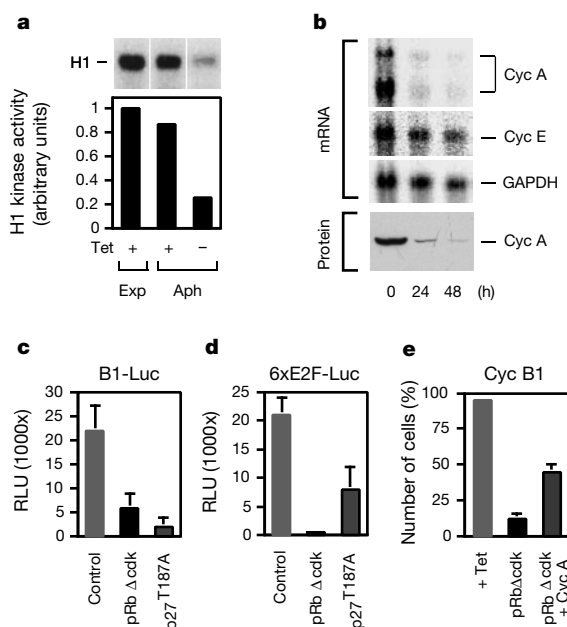


Figure 2 Cdk2 activity is essential for cyclin B1 accumulation. **a**, Cdk2-associated histone H1 kinase activity in exponentially growing (Exp) and aphidicolin-arrested (Aph) A5C1 cells, with (-Tet) or without (+Tet) induced pRb Δ cdk, quantified from a representative experiment upon identical exposure. **b**, Immunoblots of cyclin A (protein), and northern blots for cyclins (Cyc) A and E and GAPDH (mRNA) in A5C1 cells induced to express pRb Δ cdk for 24 and 48 h. **c**, Aphidicolin-treated U-2-OS cells co-injected with the B1-Luc reporter and the indicated expression plasmids, analysed *in vivo* for protein degradation as in Fig. 1b. **d**, E2F activity in live cells¹⁴ synchronized and micro-injected as in **c**, measured using the 6x2E2F-Luc reporter¹⁴. **e**, U-2-OS derived clones were synchronized by aphidicolin, induced to express the transgenes as indicated (Tet, no transgene) and immunostained for endogenous Cyclin B1.

In yeast and *Drosophila*, CDK activity is required to allow accumulation of the B-type cyclins^{11,20-23}. As induction of pRb Δ cdk in aphidicolin-arrested cells led to a fivefold reduction of Cdk2 kinase activity (Fig. 2a), we tested whether human cyclins E or A, both E2F targets¹ and activators of Cdk2, could mediate the E2F impact on cyclin B1 stability. Induction of pRb Δ cdk resulted in a marked reduction of cyclin A but not cyclin E messenger RNA levels (Fig. 2b), reflected by rapid disappearance of cyclin A (Fig. 2b) but not cyclin E^{3,5} protein, and loss of cyclin-A-dependent but only partial reduction of cyclin-E-dependent kinase activity (ref. 5; and data not shown). These effects of pRb Δ cdk were again mediated by E2F transrepression as co-expression of ectopic E2F-1 restored the cyclin A expression (data not shown). These data indicated that the reduction of Cdk2 activity following induction of pRb Δ cdk was largely due to disappearance of cyclin A. If cyclin A/Cdk2 represented an essential link between E2F-mediated transcription and stability of the B-type cyclins, direct inhibition of the Cdk2 kinase activity should mimic the effect obtained by repression of E2F. Indeed, micro-injection of a stabilized version of the Cdk2 inhibitor p27^{T187A} (ref. 24) into S-phase cells resulted in destruction of the B1-Luc protein, quantitatively even more pronounced than the effect of pRb Δ cdk (Fig. 2c). Parallel *in vivo* promoter/reporter analysis¹⁴ showed that p27^{T187A} induced destruction of cyclin B1 before it fully repressed cellular E2F activity (Fig. 2d; and data not shown), consistent with the interpretation that Cdk2 controls stability of the cyclin B1 protein downstream of E2F. A role for cyclin A was also confirmed using a reverse strategy, as co-expression of wild-type cyclin A, together with pRb Δ cdk, resulted in stabilization of cyclin B1 in a significant fraction of S-phase cells (Fig. 2e).

Mitotic cyclins are unstable in G1, and their proteolysis until G1/S transition requires association of the APC^{25,26} with the substrate-specific activator Cdh1 (Hct1 in yeast and Fizzy-related in *Drosophila*)^{11-13,21,27}. In exponentially growing U-2-OS cells, association of the APC subunit Cdc27 with Cdh1 was readily detectable, contrary to poor Cdc27-Cdh1 interaction in cells synchronized in S phase (Fig. 3a, lanes 1 and 2). Induction of pRb Δ cdk in S-phase cells markedly increased the amounts of Cdh1 within the Cdc27 immunocomplexes (Fig. 3a, lane 3), indicating that the regulation of APC-Cdh1 assembly was consistent with the differential rate of

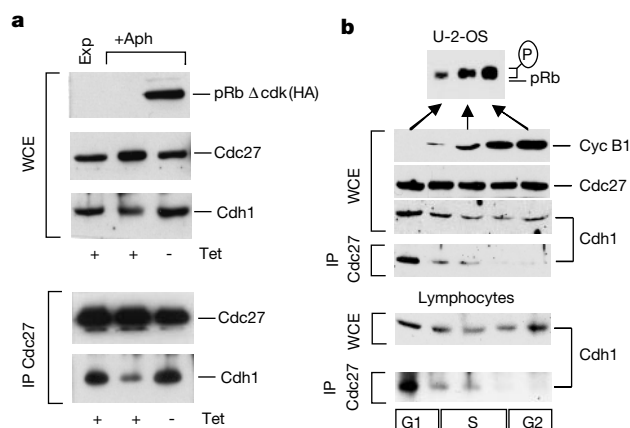


Figure 3 Reactivation of pRb in S phase promotes reassociation of APC with Cdh1. **a**, Immunoblots of Cdh1 co-immunoprecipitated with Cdc27 (IP Cdc27), compared with the abundance of the indicated proteins in whole-cell extracts (WCE) of exponentially growing (Exp) or aphidicolin-arrested (+Aph) A5C1 cells with (-Tet) or without (+Tet) pRb Δ cdk induction. **b**, Progressive loss of Cdh1 association with APC in cells beyond the G1/S transition correlates with the phosphorylation status of pRb. Immunoblots of whole-cell extracts (WCE) or Cdc27 immunoprecipitates (IP Cdc27) in matched fractions from elutriated U-2-OS cells and mitogen-stimulated lymphocytes are shown.

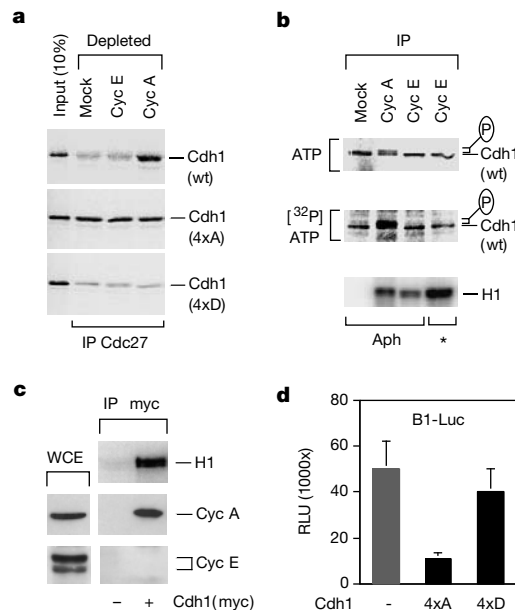


Figure 4 Phosphorylation of Cdh1 regulates Cdh1-APC assembly. **a**, *In vitro* binding of wild-type or mutant Cdh1 proteins to Cdc27 in extracts from aphidicolin-arrested U-2-OS cells, immunodepleted as indicated. **b**, Kinase assays with cyclins immunopurified from U-2-OS cells synchronized by aphidicolin (Aph) or expressing ectopic cyclin E (asterisk), using *in vitro* translated Cdh1 or histone H1 as substrates. **c**, Anti-myc immunoprecipitates from cells transiently expressing myc-tagged Cdh1, subjected to histone H1

kinase assays or immunoblotted for cyclin A or E. **d**, Aphidicolin-arrested U-2-OS cells co-injected with expression plasmids for the B1-Luc reporter and Cdh1 mutants, assayed for protein degradation *in vivo* as in Fig. 1b. Cyc, cyclin; IP, immunoprecipitates; wt, wild type.

cyclin B1 protein turnover. Parallel experiments with cells synchronized by centrifugal elutriation excluded any adverse effect of aphidicolin, and confirmed that once the U-2-OS cells passed the G1/S transition the interaction between APC and Cdh1 progressively declined, being inversely correlated with the increasing phosphorylation of the pRb family of E2F repressors, and with the gradual accumulation of cyclin B1 (Fig. 3b; and data not shown). Identical kinetics of APC-Cdh1 interactions were observed in elutriated fractions from stimulated primary human lymphocytes (Fig. 3b, lower panels), excluding the possibility that the observed phenomena were restricted to the U-2-OS model cell line. Although S-phase cells from both aphidicolin-treated or elutriated populations contained reduced levels of total Cdh1 protein, the decrease of the APC-associated Cdh1 was always more pronounced (2–4-fold, as measured by densitometric scanning). We postulated that it is mainly the APC-Cdh1 interaction that is targeted by cellular activities such as cyclin A/Cdk2, the accumulation of which requires E2F.

In support of this, wild-type Cdh1 translated *in vitro* bound poorly to the endogenous APC in extracts from S-phase cells, whereas immunodepletion of cyclin A (but not cyclin E) from S-phase extracts resulted in a significant (3–4-fold) enhancement of the wild-type Cdh1 binding to APC (Fig. 4a). Notably, a mutant of Cdh1 with four Ser/Thr residues within the putative CDK-consensus sites (conserved and regulating APC-Cdh1 assembly in yeast^{11,28}) mutated to alanine (Cdh1(4 × A)) gained the potential to associate with APC in S-phase extracts regardless of cyclin depletion (Fig. 4a). In contrast, mutating identical residues to negatively charged aspartic acid to mimic phosphorylation resulted in a Cdh1 protein that was incapable of entering APC complexes (Fig. 4a). The functional link between cyclin A and Cdh1 was further strengthened by the ability of cyclin-A-dependent, but not cyclin-E-dependent kinase, to phosphorylate Cdh1 *in vitro* (Fig. 4b), and by ectopically expressed Cdh1 associating with histone H1 kinase activity and interacting physically with cyclin A but not cyclin E

(Fig. 4c). Collectively, our data indicate that the APC-Cdh1 interaction reflects the abundance of cyclin A as a function of E2F activity, and cyclin A/Cdk2 represents an essential kinase responsible for phosphorylation of Cdh1, thereby guarding against unscheduled reactivation of APC during S phase. Consistent with this model, the ectopically expressed phosphorylation-deficient mutant Cdh1(4 × A) remained refractory to regulation *in vivo*, and resulted in destruction of the B1-Luc protein even in S-phase cells (Fig. 4d).

These results identify a new role for E2F, whose continuing activity during S phase appears essential for accumulation to physiologically effective threshold levels of mitotic regulators such as cyclin B1, a key activator of the Cdc2 kinase required to initiate cell division⁶. Critical for this mechanism is E2F-dependent expression of cyclin A, and cyclin A/Cdk2-mediated phosphorylation of Cdh1, which prevents the activatory assembly of Cdh1 with APC, thus creating an environment permissive for accumulation of APC targets. Our results help to understand the coordination of events during mammalian somatic-cell cycles, by indicating how the E2F-mediated, late-G1 commitment to replicate the genome affects also APC-dependent proteolysis and thereby completion of the cell-division cycle. Finally, as the pRb/E2F pathway is commonly targeted in oncogenesis², our data imply that the growth advantage gained through deregulating E2F in cancer cells may encompass facilitated S-phase entry as well as G2/M transition. □

Methods

Plasmids

B1-Luc reporter plasmid was constructed by cloning a fragment encoding the amino-terminal 119 amino acids of cyclin B1 upstream and in-frame with the luciferase gene within the pGL3-Control plasmid. Myc-tagged Cdh1 was generated by subcloning human Cdh1 complementary DNA into the pX^{myc} CMV-driven expression plasmid²⁹. Cdh1(4 × A) and Cdh1(4 × D) mutants were generated by substitution of serines 40, 163, 151 and threonine 121 by alanine or aspartic acid, respectively, using site-directed mutagenesis. The expression plasmid for p27^{T187A} was obtained from M. Pagano. Other plasmids are described elsewhere^{14,18,19}.

Cell culture

For induction in S phase, U-2-OS-derived A5C1 clone⁵ conditionally expressing pRbΔCdk and B5D5 clone (pRbΔCdk + Cyclin A) were presynchronized with aphidicolin (5 μg ml⁻¹) in the presence of tetracycline (1 μg ml⁻¹) for 18 h, washed and supplied with tetracycline-free medium containing aphidicolin for 24 h. Where indicated, LLNL (25 μg ml⁻¹) was added for the last 12 h before harvesting. Centrifugal elutriation of U-2-OS cells and stimulated lymphocytes were done as described³⁰.

Immunocytochemical techniques

Immunoblotting, immunoprecipitation, immunostaining and *in vitro* kinase assays are published^{14,29}, as is the rabbit antiserum against Cdc27¹³; the Cdh1 rabbit antiserum was raised and characterized by E.K. and J.-M.P.

In vitro binding assays

Cells synchronized by aphidicolin were lysed in hypotonic buffer (25 mM Hepes pH 7.5, 5 mM KCl, 1.5 mM MgCl₂, 1 mM DTT); 200 μg lysate was combined with 1 μl *in vitro* translated Cdh1 labelled with [³⁵S]methionine, 1 mM ATP, and incubated in a total volume of 20 μl for 1 h at 30 °C. Cdh1 co-immunoprecipitated with Cdc27 antibody was quantified by phosphorimager (Molecular Dynamics).

In vivo luminescence imaging

Photon emission by live cells microinjected with cyclin B1-Luc or 6x2F-Luc reporter constructs, together with expression plasmids specified in figure legends, was measured as described¹⁴.

Received 9 June; accepted 19 August 1999.

- Dyson, N. The regulation of E2F by pRB-family proteins. *Genes Dev.* **12**, 2245–2262 (1998).
- Weinberg, R. A. The retinoblastoma protein and cell cycle control. *Cell* **81**, 323–330 (1995).
- Knudsen, E. S., Buckmaster, C., Chen, T. T., Feramisco, J. R. & Wang, J. Y. Inhibition of DNA synthesis by RB: effects on G1/S transition and S-phase progression. *Genes Dev.* **12**, 2278–2292 (1998).
- Chew, Y. P., Ellis, M., Wilkie, S. & Mittnacht, S. pRB phosphorylation mutants reveal role of pRB in regulating S phase completion by a mechanism independent of E2F. *Oncogene* **17**, 2177–2186 (1998).
- Lukas, J., Sørensen, C. S., Lukas, C., Santoni-Rugui, E. & Bartek, J. p16INK4A, but not constitutively active pRB, can impose a sustained G1 arrest: molecular mechanism and implications for oncogenesis. *Oncogene* **18**, 3930–3935 (1999).
- Nurse, P. Universal control mechanism regulating onset of M-phase. *Nature* **344**, 503–508 (1990).
- Pines, J. & Hunter, T. Human cyclins A and B1 are differentially located in the cell and undergo cell cycle-dependent nuclear transport. *J. Cell Biol.* **115**, 1–17 (1991).
- Townsend, F. M. & Ruderman, J. V. Proteolytic ratchets that control progression through mitosis. *Trends Cell Biol.* **8**, 238–244 (1998).
- King, R. W., Deshaies, R. J., Peters, J. M. & Kirschner, M. W. How proteolysis drives the cell cycle. *Science* **274**, 1652–1659 (1996).
- Visintin, R., Prinz, S. & Amon, A. CDC20 and CDH1: a family of substrate-specific activators of APC-dependent proteolysis. *Science* **278**, 460–463 (1997).
- Zachariae, W., Schwab, M., Nasmyth, K. & Seufert, W. Control of cyclin ubiquitination by CDK-regulated binding of Hct1 to the anaphase promoting complex. *Science* **282**, 1721–1724 (1998).
- Fang, G., Yu, H. & Kirschner, M. W. Direct binding of CDC20 protein family members activates the anaphase-promoting complex in mitosis and G1. *Mol. Cell* **2**, 163–171 (1998).
- Kramer, E. R., Gieffers, C., Holz, G., Hengstschlager, M. & Peters, J. M. Activation of the human anaphase-promoting complex by proteins of the CDC20/Fizzy family. *Curr. Biol.* **8**, 1207–1210 (1998).
- Lukas, J. *et al.* Cyclin E-induced S phase without activation of the pRb/E2F pathway. *Genes Dev.* **11**, 1479–1492 (1997).
- Solomon, M. J., Glotzer, M., Lee, T. H., Philippe, M. & Kirschner, M. W. Cyclin activation of p34cdc2. *Cell* **63**, 1013–1024 (1990).
- Glotzer, M., Murray, A. W. & Kirschner, M. W. Cyclin is degraded by the ubiquitin pathway. *Nature* **349**, 132–138 (1991).
- Brandeis, M. & Hunt, T. The proteolysis of mitotic cyclins in mammalian cells persists from the end of mitosis until the onset of S phase. *EMBO J.* **15**, 5280–5289 (1996).
- Johnson, D. G., Cress, W. D., Jakoi, L. & Nevins, J. R. Oncogenic capacity of the E2F1 gene. *Proc. Natl. Acad. Sci. USA* **91**, 12823–12827 (1994).
- Wu, C.-L., Classon, M., Dyson, N. & Harlow, E. Expression of dominant-negative mutant DP-1 blocks cell cycle progression in G1. *Mol. Cell Biol.* **16**, 3698–3706 (1994).
- Amon, A., Irniger, S. & Nasmyth, K. Closing the cell cycle circle in yeast: G2 cyclin proteolysis initiated at mitosis persists until the activation of G1 cyclins in the next cycle. *Cell* **77**, 1037–1050 (1994).
- Sigrist, S. J. & Lehner, C. F. *Drosophila* fizzy-related down-regulates mitotic cyclins and is required for cell proliferation arrest and entry into endocycles. *Cell* **90**, 671–681 (1997).
- Knoblich, J. A. *et al.* Cyclin E controls S phase progression and its down-regulation during *Drosophila* embryogenesis is required for the arrest of cell proliferation. *Cell* **77**, 107–120 (1994).
- Irniger, S. & Nasmyth, K. The anaphase-promoting complex is required in G1 arrested yeast cells to inhibit B-type cyclin accumulation and to prevent uncontrolled entry into S-phase. *J. Cell Sci.* **110**, 1523–1531 (1997).
- Nguyen, H., Gitig, D. M. & Koff, A. Cell-free degradation of p27^{kip1}, a G1 cyclin-dependent kinase inhibitor, is dependent on CDK2 activity and the proteasome. *Mol. Cell Biol.* **19**, 1190–1201 (1999).
- Sudakin, V. *et al.* The cyclosome, a large complex containing cyclin-selective ubiquitin ligase activity, targets cyclins for destruction at the end of mitosis. *Mol. Biol. Cell* **6**, 185–197 (1995).
- King, R. W. *et al.* A 20S complex containing CDC27 and CDC16 catalyzes the mitosis-specific conjugation of ubiquitin to cyclin B. *Cell* **81**, 279–288 (1995).
- Schwab, M., Lutum, A. S. & Seufert, W. Yeast Hct1 is a regulator of Clb2 cyclin proteolysis. *Cell* **90**, 683–693 (1997).
- Jaspersen, S., Charles, J. F. & Morgan, D. O. Inhibitory phosphorylation of the APC regulator Hct1 is controlled by the kinase Cdc28 and the phosphatase Cdc14. *Curr. Biol.* **9**, 227–236 (1999).

- Lukas, J. *et al.* Retinoblastoma-protein-dependent cell-cycle inhibition by the tumour suppressor p16. *Nature* **375**, 503–506 (1995).
- Lukas, J. *et al.* Cyclin D2 is a moderately oscillating nucleoprotein required for G1 phase progression in specific cell types. *Oncogene* **10**, 2125–2134 (1995).

Acknowledgements

We dedicated this letter to the memory of M. Strauss. We thank C. Gieffers, K. Helin, M. Pagano and C.-L. Wu for providing important reagents, M. Rohde for computer assistance, K. Hansen for help with T-lymphocyte stimulation, and the Danish Cancer Society, the Human Frontier Science Programme and the Danish Medical Research Council for grant support.

Correspondence and requests for materials should be addressed to J.B. (e-mail: bartek@biobase.dk).

NMR structure and mutagenesis of the inhibitor-of-apoptosis protein XIAP

Chaohong Sun*, Mengli Cai*, Angelo H. Gunasekera, Robert P. Meadows, Hong Wang, Jun Chen, Haichao Zhang, Wei Wu, Nan Xu, Shi-Chung Ng & Stephen W. Fesik

Pharmaceutical Discovery Division, Abbott Laboratories, Abbott Park, Illinois 60064, USA

* These authors contributed equally to this work

The inhibitor-of-apoptosis (IAP) family of proteins, originally identified in baculoviruses¹, regulate programmed cell death in a variety of organisms^{2–6}. IAPs inhibit specific enzymes (caspases) in the death cascade^{7–11} and contain one to three modules of a common 70-amino-acid motif called the BIR domain¹². Here we describe the nuclear magnetic resonance structure of a region encompassing the second BIR domain (BIR2) of a human IAP family member, XIAP (also called hIAP or MIHA). The structure of the BIR domain consists of a three-stranded antiparallel β-sheet and four α-helices and resembles a classical zinc finger¹³. Unexpectedly, conserved amino acids within the linker region between the BIR1 and BIR2 domains were found to be critical for inhibiting caspase-3. The absence or presence of these residues may explain the differences in caspase inhibition observed for different truncated and full-length IAPs^{10,11}. Our data further indicate that these residues may bind to the active site and that the BIR domain may interact with an adjacent site on the enzyme.

A region of XIAP has been identified that contains the second BIR domain (BIR2), which inhibits caspases 3 and 7 with inhibition constants of 2–5 nM¹⁰. NMR spectra of this portion of XIAP (residues 124–260) yielded broad signals indicative of aggregation. To obtain a suitable NMR sample, we truncated the protein and prepared site-directed mutants. Truncation of 24 residues from the amino terminus resulted in a total loss of caspase inhibition and only a marginal improvement in the NMR signals; however, marked improvements in the quality of the NMR spectra were obtained by removing 20 residues from the carboxy terminus and mutating two nonconserved cysteines (C202A and C213G) to residues located at these positions in other IAPs⁷. This truncated mutant protein has a similar inhibitory potency to that of the wild-type protein (residues 124–260) (Table 1) and was suitable for structure determination by NMR.

The structure of (C202A/C213G)XIAP (residues 124–240) was determined from a total of 1,092 NMR-derived restraints. As shown in the ensemble of NMR structures (Fig. 1a), the backbone of the BIR2 domain (residues 163–230) and a few residues outside this region are well defined by the NMR data. In contrast, the structures

of 29 residues at the N terminus and 9 residues at the C terminus of our construct could not be characterized owing to the paucity of nuclear Overhauser effects (NOEs). The small heteronuclear NOEs (<0.5) and small ^1H - ^1H residual dipolar couplings (<1 Hz) of the N- and C-terminal residues of XIAP measured in the presence of phage^{14,15} indicate that this portion of the protein may be flexible and may tumble independently of the BIR2 domain (data not shown). Despite its flexibility, however, the N terminus contains a short α -helix (residues 137–140) and may transiently interact with the BIR2 domain, as indicated by small chemical-shift changes in BIR2 upon mutation of the N-terminal residues.

The structure of the BIR2 domain of XIAP resembles a classical zinc finger¹³ and consists of a three-stranded antiparallel β -sheet and four α -helices (Fig. 1b). Amino acids C200, C203, H220 and C227, which are present in all IAPs⁷, chelate zinc and stabilize the overall fold. Additional stabilization is achieved by hydrophobic interactions within the core of the protein between the conserved residues F170, W173, L179, L184, L189, V198 and W210. The prolines and glycines that are conserved (Fig. 1c) are located in turns. Notably, G188, which is found in all IAPs, is located in a sharp turn between the third α -helix and the first strand of the β -sheet.

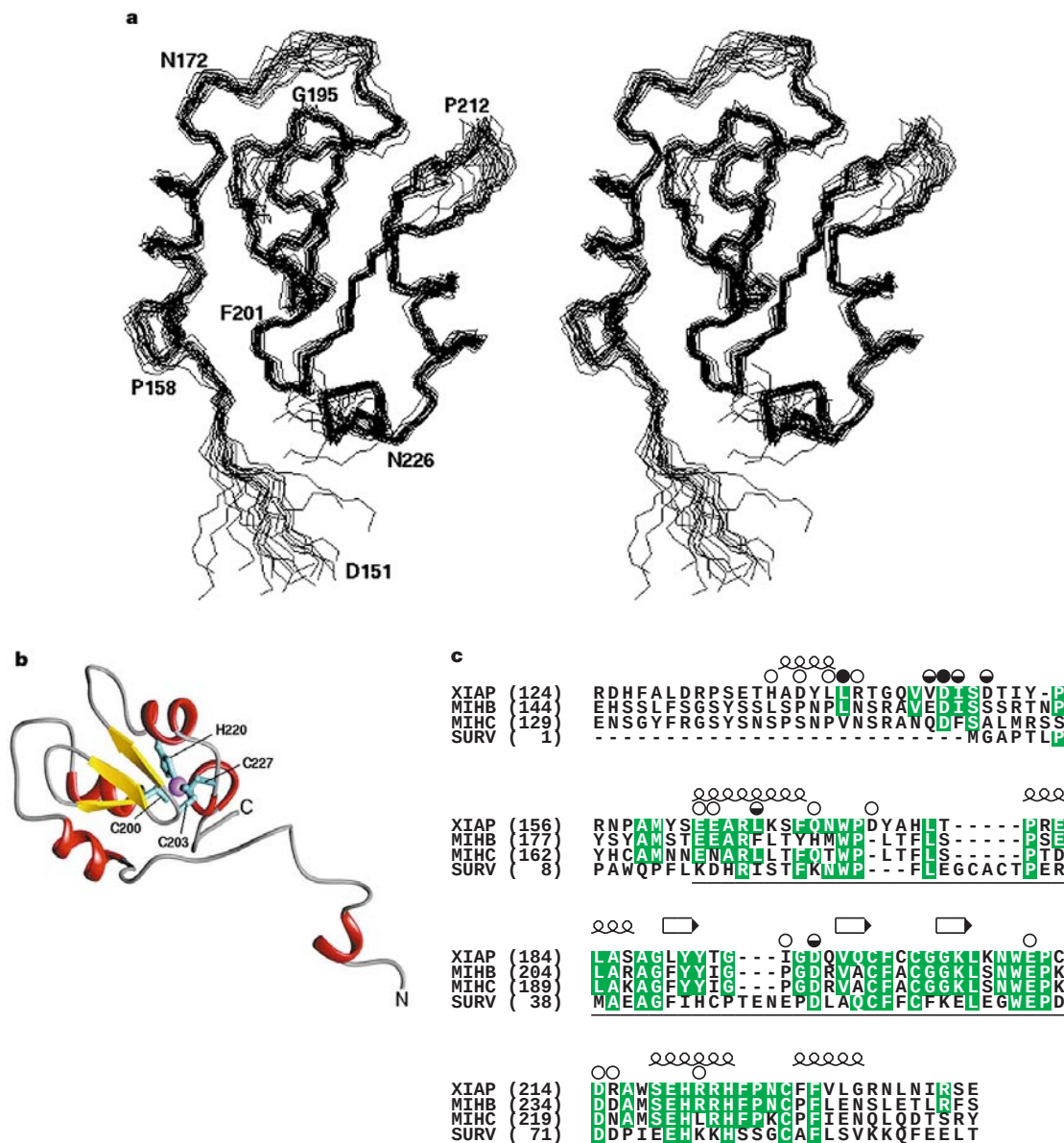


Figure 1 Structure of the BIR2 domain of XIAP. **a**, Stereoview of 20 superimposed NMR-derived structures of (C202A/C213G)XIAP (residues 151–231). The relative mean standard deviation (r.m.s.d.) of the mean coordinate positions for residues 153–231 is 0.70 ± 0.08 for the backbone atoms (N, C α , C') and 1.20 ± 0.10 for all heavy atoms. No distance restraint is violated by more than $0.4 \text{ \AA}$ in any of the structures, and no torsional restraint is violated by more than 5° . For the ensemble of final structures, the X-PLOR NOE potential energy is $17.6 \pm 6.8 \text{ kcal mol}^{-1}$, the torsional energy is $0.78 \pm 0.35 \text{ kcal mol}^{-1}$ and the van der Waals energy is $-542 \pm 15 \text{ kcal mol}^{-1}$. **b**, Ribbons³⁰ representation of the averaged minimized NMR structure of (C202A/

C213G)XIAP (residues 124–240). The side chains of the residues that chelate zinc are shown. **c**, Sequence alignment of the BIR2 domains and surrounding residues of human IAPs that inhibit caspases (GenBank accession number: hIAP/XIAP, U32974; MIHB, U37547; MIHC, U37546; SURV, U75285). Identical residues are highlighted in green. The secondary structural elements in the NMR structure of XIAP are shown above the sequence. Circles denote residues mutated to alanine. Open circles denote no significant change in caspase inhibition for the mutant protein; half-filled circles denote more than a 3-fold decrease; and filled circles denote more than a 10-fold decrease in the inhibition of caspase-3. The amino acids corresponding to the BIR2 domain¹⁰ are underlined.

The surface of the protein (Fig. 2a) consists mostly of charged residues. On one side of the protein, there is a negatively charged region which is composed of E211, D214 and E219 (Fig. 2a, left). On the other side (Fig. 2a, right), there is a cluster of positively charged residues (R215, R221, R222 and R233). The side chains of a few hydrophobic residues appear on the surface of the protein; however, these residues are not conserved within the IAP family.

To determine which amino acids of XIAP are responsible for inhibiting caspases, site-directed mutant proteins were prepared and tested for their ability to inhibit caspase-3 (Table 1). Because of the importance of acidic residues in substrates and inhibitors of caspases^{16,17}, our initial studies focused on mutagenesis of the aspartic and glutamic amino acids. Mutation of the acidic residues (E211A and D214A) in the negatively charged region of the protein or any of the other acidic residues within the BIR2 domain (except for D196), however, did not alter the ability of the mutant proteins to inhibit caspase-3. Similarly, mutations of positively charged residues (R215 and R221) and hydrophobic residues (L167 and I194) on the surface of the BIR2 domain had little or no effect on caspase inhibition (Table 1). In contrast, several residues N-terminal to the BIR2 domain were found to be critical for inhibiting caspase-3. An attenuation of caspase-3 inhibition was observed for alanine mutants of L141, V147, I149 and D151 (Table 1). Moreover, a complete loss of inhibitory potency was observed for the D148A mutant. These results are consistent with the inability of N-terminally truncated proteins to inhibit caspase-3 even though they adopt the wild-type fold (data not shown). Although very little sequence similarity is generally found outside the BIR domains, L141, D148 and I149 are conserved in the linker region between BIR1 and BIR2 in human IAPs that potently inhibit caspases (Fig. 1c). Another human IAP, survivin, does not contain this region (Fig. 1c), and would not be predicted to potently inhibit caspases. This is consistent with co-immunoprecipitation experiments indicating that survivin does not bind to caspases as well as does XIAP¹¹. Our data also explain the lack of caspase inhibition observed for a BIR1-containing protein that does not possess this linker region¹⁰. Thus, the differences in caspase inhibition observed for the individual BIR domains may be due to the presence or

absence of these residues. We confirmed this by converting the BIR1 domain of XIAP into a potent inhibitor of caspase-3 by attaching this linker region to either the N or C terminus (Table 1); however, a peptide corresponding to the linker region was insufficient by itself to inhibit caspase-3 (Table 1), indicating that the BIR domain may also be important for caspase inhibition.

To explore the molecular mechanism by which BIR-domain-containing proteins inhibit caspases, we conducted NMR studies on the interaction of isotopically labelled (C202A/C213G) XIAP (residues 124–240) with both a caspase-8-cleaved and a pro-form of a caspase-3 mutant (C285A). This enzymatically inactive caspase-3 mutant was chosen because wild-type caspase-3 cleaved XIAP under the conditions of the NMR experiments. Upon binding to the C285A caspase-3-mutant that was cleaved to mimic its activated form, amide chemical-shift changes were observed for residues in both the linker region and the BIR2 domain (Fig. 2b). In contrast, chemical-shift changes were only observed for the BIR2 domain upon binding to the pro-form of the C285A caspase-3 mutant (Fig. 2c). Furthermore, the binding of XIAP to the pro-form, as measured by NMR, was much weaker (dissociation constant, $k_d = 30 \mu\text{M}$) compared with the activated form of the enzyme. These results indicate that the linker region of XIAP may bind to caspase-3 in the active site, but that it cannot interact with the pro-form of the enzyme in which the active site has a different conformation or is sterically occluded. This is supported by the similarity of the residues in the linker region (-DISD-) to caspase substrates/inhibitors, and by the marked reduction in binding to activated caspase-3 in the presence of a tetrapeptide inhibitor (Fig. 3).

IAPs are important regulators of apoptosis in many organisms and inhibit specific caspases in cell-death signalling pathways. The region of IAPs responsible for mediating programmed cell death and caspase inhibition contains the BIR domains. Although BIR domains are highly conserved, our structure of the BIR2 domain of XIAP differs from the structure of the BIR3 domain of MIHB/c-IAP-1 that lacks a β -sheet¹⁸. BIR domains also display differences in their ability to inhibit caspases¹⁰; however, the differences in caspase inhibition may not be related to the individual BIR domains but may reflect differences in the surrounding amino acids. Indeed, in

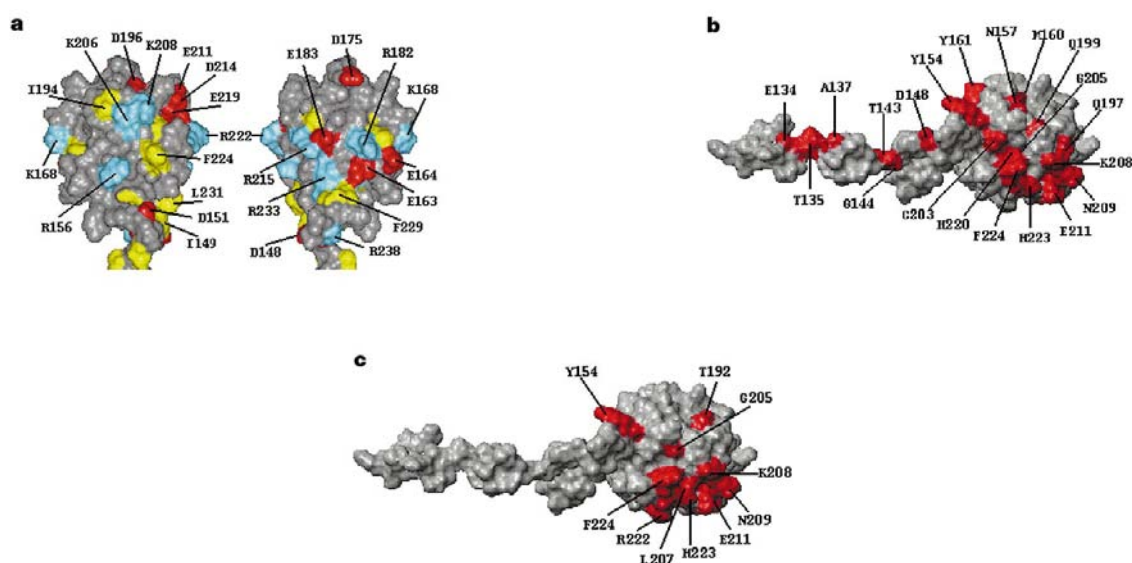


Figure 2 Surface of the BIR2 domain of XIAP. **a**, Two views of the surface of the averaged minimized NMR structure of (C202A/C213G)XIAP (residues 147–240). Lys and Arg side-chain atoms are blue; Asp and Glu side-chain atoms are red; Phe, Val, Leu, and Ile side-chain atoms are yellow. **b**, **c**, Surface representations of the averaged minimized NMR

structure of (C202A/C213G)XIAP (residues 124–240). The residues with well resolved amide resonances and significant chemical-shift changes upon binding to the caspase-8-cleaved (**b**) or pro-form (**c**) of C285A mutant caspase-3 are red.

Table 1 Caspase-3 inhibition (IC₅₀s) of XIAP mutant proteins

Mutant	IC ₅₀ (nM)	Mutant	IC ₅₀ (nM)
XIAP(124–260)*	11 ± 2	E163A	31 ± 2
XIAP(124–240)	14 ± 7	E164A	24 ± 2
H136A	17 ± 2	L167A	86 ± 15
D138A	8 ± 1	Q171A	8 ± 3
L140A	42 ± 2	D175A	11 ± 3
L141A	30% @ 1 μM	I194A	11 ± 2
R142A	47 ± 6	D196A	200†
V147A	162 ± 21	E211A	16 ± 3
D148A	0% @ 1 μM	D214A	23 ± 3
I149A	139 ± 55	R215A	11 ± 1
D151A	61 ± 18	R221A	3 ± 2
Linker(134–154)	0% @ 1 μM	Bir1(1–123)	0% @ 1 μM
Linker-Bir1 (124–159 + 23–123)	71 ± 15	Bir1-Linker (1–168)	67 ± 18

IC₅₀s are reported as the mean ± s.d. of at least three independent measurements. ¹⁵N-¹H correlation spectra were recorded for uniformly ¹⁵N-labelled samples of all the mutant proteins that displayed a decrease in caspase-3 inhibition. For all of these proteins, only a few chemical shift differences were observed compared with the wild-type protein, indicating that the overall fold was not disrupted and was not the cause of the decreased caspase inhibition.

*In addition to the listed mutations, all the mutant proteins contain the C202A/C213G double mutation except for XIAP(124–260).

†This mutant protein exhibited a time-dependent loss of caspases-3 inhibition which could be attributed to its instability.

XIAP, residues N-terminal to the BIR2 domain are required for caspase inhibition. On the basis of our NMR studies, these residues may bind to the active site of caspase-3 while the BIR domain binds to an adjacent site, resulting in a high-affinity complex.

Methods

NMR sample preparation

XIAP (residues 124–260) was cloned from a Jurkat complementary DNA library. From this template, (C202A/C213G)XIAP (residues 124–240) was engineered and expressed in *Escherichia coli* using the pET15b vector (Novagen). For the NMR experiments, uniformly ¹⁵N- and ¹³C-labelled proteins were prepared by growing the *E. coli* strain BL21(DE3) in minimal medium containing a trace amount of zinc and ¹⁵NH₄Cl with or without [U-¹³C]glucose. Recombinant proteins were purified by affinity chromatography on a nickel-IDA column (Invitrogen) and gel filtration (Superdex 200, Pharmacia). The N-terminal His tag was cleaved with thrombin leaving four extra residues (GSHM) at the N terminus. NMR samples contained 0.6 mM protein in 50 mM Tris pH 7.5, 300 mM KCl, 50 μM Zn(Ac)₂ and 0.5 mM DTT in 95% H₂O/5% ²H₂O or 100% ²H₂O.

Structure determination

NMR spectra were acquired at 30 °C on Bruker 500, 600 or 800 MHz NMR spectrometers. The ¹H, ¹³C and ¹⁵N resonances of the backbone were assigned using triple resonance experiments (CBCANH and CBCA(CO)NH), and the ¹³C and ¹H resonances of the side chains were assigned from 3D HCCH-total correlation spectroscopy (TOCSY), HBHA(CBCACO)NH, and (H)C(CO)NH-TOCSY experiments¹⁹. Stereospecific assignment of methyl groups of valine and leucine residues were obtained using a fractionally ¹³C-labelled sample as described²⁰. Distance restraints were obtained from ¹⁵N- or ¹³C-resolved 3D nuclear Overhauser enhancement spectroscopy (NOESY) experiments. ϕ and χ_1 angle restraints were derived from coupling constants measured using described NMR

experiments^{21,22}. In addition, ϕ and ψ torsional restraints were obtained from an analysis of the C', N, C α , H α and C β chemical shifts using the TALOS program²³. Slowly exchanging amide protons were identified from a series of ¹H-¹⁵N HSQC spectra recorded after the H₂O buffer was changed to a ²H₂O buffer. Structures of XIAP were calculated with torsional angle dynamics²⁴ and simulated annealing²⁵ using a modified version of the X-PLOR program²⁶. The structure calculations employed 796 inter-residue proton-proton distance restraints, 38 hydrogen-bond-distance restraints, 38 ϕ , 31 ψ , 47 χ_1 and 7 χ_2 angle restraints, and 65 N-H, 33 N-C' and 37 H^N-C' residual dipolar couplings. The residual dipolar couplings measured in the presence of Pfl phage (15 mg ml⁻¹) were included in the structure refinement as described²⁷. The NOE-derived restraints were categorized into 4 bins (2.7, 3.5, 5.0 and 6.0 Å) for NOEs involving ¹³C-attached protons and 5 bins (2.9, 3.3, 5.0, 6.0 and 7.0 Å) for NOEs involving ¹⁵N-attached protons. An additional 0.5 Å was added for NOEs involving methyl groups.

NMR-based binding experiments

¹H-¹⁵N amide chemical shift changes of [¹⁵N,²H](C202A/C213G)XIAP (residues 124–240) were monitored upon the addition of the cleaved or pro-form of a C285A caspase-3 mutant using a sensitivity enhanced version²⁸ of the TROSY experiment²⁹. The C285A caspase-3 mutant was cleaved by the addition of caspase-8 (Pharmingen).

Site-directed mutagenesis

Mutant proteins were prepared using the Quick-change site-directed mutagenesis kit (Stratagene). The coding region was sequenced in each case.

Caspase-3 inhibition assay

The enzymatic activity of caspase-3 was monitored at 37 °C in a reaction mixture containing 2.4 ng of purified human recombinant caspase-3 and 100 μM DEVD-AMC in a buffer composed of 20 mM Hepes (pH 7.7), 100 mM NaCl, 0.1% CHAPS and 10 mM DTT. Concentrations for 50% inhibition (IC₅₀s) were determined from the enzymatic activities measured 40 min after initiating the reaction by the addition of DEVD-AMC in different concentrations of XIAP proteins.

Caspase-3 binding assay

(C202A/C213G)XIAP (residues 124–240) (1 μg) and activated His-tagged caspase-3 (2 μg) were incubated for 1 h at 4 °C with 5 μl of Ni²⁺-Sephacrose in the presence and absence of 10 μM DEVD-aldehyde or a control peptide GQVGRQAAIGDDINR. The buffer (100 μl) used in the binding experiments contained 50 mM Tris pH 7.5, 100 mM KCl, 10% sucrose, 0.1% CHAPS and 5 mM dithiothreitol. The Ni²⁺ beads were washed five times with the same buffer before boiling in Laemmli buffer and analysis using a 16.4% Tris/Tricine/SDS gel. The proteins were visualized using coomassie blue stain.

Received 25 June; accepted 13 August 1999.

- Crook, N. E. Clem, R. J. & Miller, L. K. An apoptosis-inhibiting baculovirus gene with a zinc finger-like motif. *J. Virol.* **67**, 2168–2174 (1993).
- Hay, B. A., Wassarman, D. A. & Rubin, G. M. Drosophila homologs of baculovirus inhibitor of apoptosis proteins function to block cell death. *Cell* **83**, 1253–1262 (1995).
- Roy, N. *et al.* The gene for neuronal apoptosis inhibitory protein is partially deleted in individuals with spinal muscular atrophy. *Cell* **80**, 167–178 (1995).
- Duckett, C. S. *et al.* A conserved family of cellular genes related to the baculovirus iap gene and encoding apoptosis inhibitors. *EMBO J.* **15**, 2685–2694 (1996).
- Uren, A. G., Pakusch, M., Hawkins, C. J., Puls, K. L. & Vaux, D. L. Cloning and expression of apoptosis inhibitory protein homologs that function to inhibit apoptosis and/or bind tumor necrosis factor receptor-associated factors. *Proc. Natl Acad. Sci. USA* **93**, 4974–4978 (1996).
- Ambrosini, G., Adida, C. & Altieri, D. C. A novel anti-apoptosis gene, survivin, expressed in cancer and lymphoma. *Nature Med.* **3**, 917–921 (1997).
- Hawkins, C. J., Uren, A. G., Hacker, G., Medcalf, R. L. & Vaux, D. L. Inhibition of interleukin 1 β -converting enzyme-mediated apoptosis of mammalian cells by baculovirus IAP. *Proc. Natl Acad. Sci. USA* **93**, 13786–13790 (1996).
- Deveraux, Q. L., Takahashi, R., Salvesen, G. S. & Reed, J. C. X-linked IAP is a direct inhibitor of cell-death proteases. *Nature* **388**, 300–304 (1997).
- Roy, N., Deveraux, Q. L., Takahashi, R., Salvesen, G. S. & Reed, J. C. The c-IAP-1 and c-IAP-2 proteins are direct inhibitors of specific caspases. *EMBO J.* **16**, 6914–6925 (1997).
- Takahashi, R. *et al.* A single BIR domain of XIAP sufficient for inhibiting caspases. *J. Biol. Chem.* **273**, 7787–7790 (1998).
- Tamm, I. *et al.* IAP-family protein survivin inhibits caspase activity and apoptosis induced by Fas(CD95), Bax, caspases, and anti cancer drugs. *Cancer Res.* **58**, 5315–5320 (1998).
- Uren, A. G., Coulsen, E. J. & Vaux, D. L. Conservation of baculovirus inhibitor of apoptosis repeat proteins (BIRPs) in viruses, nematodes, vertebrates and yeasts. *Trends Biochem. Sci.* **23**, 159–162 (1998).
- Schwabe, J. W. R. & Klug, A. Zinc mining for protein domains. *Nature Struct. Biol.* **1**, 345–349 (1994).
- Hansen, M. R., Mueller, L. & Pardi, A. Tunable alignment of macromolecules by filamentous phage yields dipolar coupling interactions. *Nature Struct. Biol.* **5**, 1065–1074 (1998).
- Clare, G. M., Starich, M. R. & Gronenborn, A. M. Measurement of residual dipolar couplings of macromolecules aligned in the nematic phase of a colloidal suspension of rod-shaped viruses. *J. Am. Chem. Soc.* **120**, 10571–10572 (1998).
- Thornberry, N. A. *et al.* A combinatorial approach defines specificities of members of the caspase family and granzyme B. *J. Biol. Chem.* **272**, 17907–17911 (1997).
- Rotunda, J. *et al.* The three-dimensional structure of apopain/CPP32, a key mediator of apoptosis. *Nature Struct. Biol.* **3**, 619–625 (1996).
- Hinds, M. G., Norton, R. S., Vaux, D. L. & Day, C. L. Solution structure of a baculoviral inhibitor of apoptosis (IAP) repeat. *Nature Struct. Biol.* **6**, 648–651 (1999).

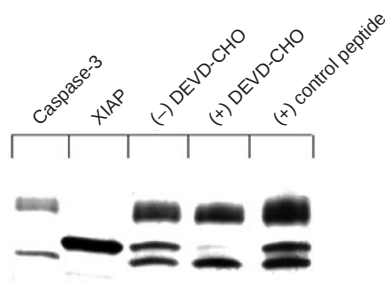


Figure 3 Binding of (C202A/C213G)XIAP (residues 124–240) to His-tagged caspase-3 in the presence and absence of the caspase-3 inhibitor DEVD aldehyde. The band corresponding to XIAP in the presence (+) of the DEVD aldehyde (DEVD-CHO) was reduced to 20% ± 3% (*n* = 3) of that in the absence (–) of the caspase-3 inhibitor. Far right lane, a control peptide (GQVGRQAAIGDDINR) that does not bind to caspase-3.

19. Clore, G. M. & Gronenborn, A. M. Multidimensional heteronuclear nuclear magnetic resonance of proteins. *Methods Enzymol.* **239**, 349–363 (1994).
20. Neri, D., Szyperski, T., Otting, G., Senn, H. & Wüthrich, K. Stereospecific nuclear magnetic resonance assignments of the methyl groups of valine and leucine in the DNA-binding domain of the 434 repressor by biosynthetically directed fractional ^{13}C labeling. *Biochemistry* **28**, 7510–7516 (1989).
21. Vuister, G. W. & Bax, A. Quantitative J correlation: A new approach for measuring homonuclear three-bond J ($^3\text{H}^3\text{H}$) coupling constants in ^{15}N -enriched proteins. *J. Am. Chem. Soc.* **115**, 7772–7777 (1993).
22. Hu, J.-S., Grzesiek, S. & Bax, A. Two-dimensional NMR methods for determining χ_1 angles of aromatic residues in proteins from three-bond J_{CH} and J_{NC} couplings. *J. Am. Chem. Soc.* **119**, 1803–1804 (1997).
23. Cornilescu, G., Delaglio, F. & Bax, A. Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J. Biomol. NMR* **13**, 289–302 (1999).
24. Stein, E. G., Rice, L. M. & Brünger, A. T. Torsion angle molecular dynamics: a new efficient tool for NMR structure calculation. *J. Magn. Reson. Ser. B* **124**, 154–164 (1997).
25. Nilges, M., Gronenborn, A. M., Brünger, A. T. & Clore, G. M. Determination of three-dimensional structures of proteins by simulated annealing with interproton distance restraints: application to crambin, potato carboxypeptidase inhibitor and barley serine proteinase inhibitor 2. *Protein Eng.* **2**, 27–38 (1988).
26. Brünger, A. T. *X-PLOR 3.1 Manual* (Yale Univ. Press, New Haven, 1992).
27. Tjandra, N., Omichinski, J. G., Gronenborn, A. M., Clore, G. M. & Bax, A. Use of dipolar ^1H – ^{15}N and ^1H – ^{13}C couplings in the structure determination of magnetically oriented macromolecules in solution. *Nature Struct. Biol.* **4**, 732–738 (1997).
28. Rance, M., Loria, J. P. & Palmer, A. G. III Sensitivity improvement of transverse relaxation-optimized spectroscopy. *J. Magn. Res.* **136**, 92–101 (1999).
29. Pervushin, K., Riek, R., Wider, G. & Wüthrich, K. Attenuated T_2 relaxation by mutual cancellation of dipole–dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proc. Natl Acad. Sci. USA* **94**, 12366–12371 (1997).
30. Carson, M. J. Ribbon models of macromolecules. *J. Mol. Graphics* **5**, 103–106 (1987).

Acknowledgements

We thank R. Mendoza for technical assistance; A. Pardi for the starter cultures of *Pseudomonas aeruginosa* and Pfl phage; M. Clore for a modified version of the X-PLOR program; F. Delaglio and A. Bax for the TALOS program; and J. Wu from IDUN Pharmaceuticals for the caspase-3 clone.

Correspondence and requests for materials should be addressed to S.W.F. (e-mail: fesiks@pprd.abbott.com). Coordinates for the NMR structure of (C202A/C213G)XIAP (residues 124–240) have been deposited in the Brookhaven Protein Data Bank under the accession code 1C9Q.

High-resolution X-ray structure of an early intermediate in the bacteriorhodopsin photocycle

Karl Edman^{*,†}, Peter Nollert^{†,§}, Antoine Royant^{†,||}, Hassan Belrhali[¶], Eva Pebay-Peyroula^{||}, Janos Hajdu^{*}, Richard Neutze^{*} & Ehud M. Landau[§]

^{*} Department of Biochemistry, Uppsala University, Biomedical Centre, Box 576, S-751 23 Uppsala, Sweden

[§] Department of Molecular Microbiology, Biozentrum, University of Basel, Klingelbergstrasse 70, CH-4056 Basel, Switzerland

^{||} Institut de Biologie Structurale, CEA-CNRS-Université Joseph Fourier, 41 rue Jules Horowitz, F-38027 Grenoble Cedex 1, France

[¶] European Synchrotron Radiation Facility, 6 rue Jules Horowitz, BP 220, F-38043 Grenoble Cedex, France

[†] These authors contributed equally to this work

Bacteriorhodopsin is the simplest known photon-driven proton pump¹ and as such provides a model for the study of a basic function in bioenergetics. Its seven transmembrane helices² encompass a proton translocation pathway containing the chromophore, a retinal molecule covalently bound to lysine 216 through a protonated Schiff base, and a series of proton donors and acceptors. Photoisomerization of the all-*trans* retinal to the 13-*cis* configuration initiates the vectorial translocation of a

proton from the Schiff base, the primary proton donor, to the extracellular side, followed by reprotonation of the Schiff base from the cytoplasm. Here we describe the high-resolution X-ray structure of an early intermediate in the photocycle of bacteriorhodopsin, which is formed directly after photoexcitation. A key water molecule is dislocated, allowing the primary proton acceptor, Asp 85, to move. Movement of the main-chain Lys 216 locally disrupts the hydrogen-bonding network of helix G, facilitating structural changes later in the photocycle.

During its photocycle, bacteriorhodopsin passes through a series of structural intermediates with well-defined lifetimes and spectral properties (Fig. 1a). The primary photochemical reaction is the isomerization of retinal (Fig. 1b), from which a proton's climb against the transmembrane electrostatic potential of the purple membrane follows. To elucidate the initial structural changes coupled to the proton-pumping mechanism, we trapped at low temperature an early intermediate of the photocycle within wild-type bacteriorhodopsin crystals grown in a lipidic cubic phase^{3,4} and determined its X-ray structure. Early intermediates can build up within crystals at low temperatures when there is insufficient thermal energy to traverse barriers associated with the formation of later states^{5,6}.

We used single-crystal microspectrophotometry⁷ to characterize

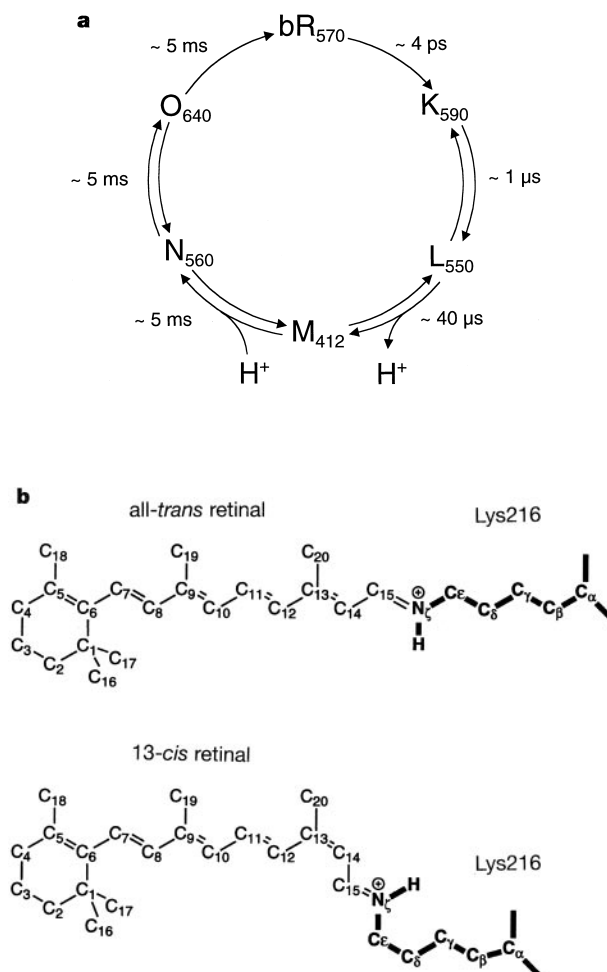


Figure 1 Key events in the photocycle of bacteriorhodopsin (bR). **a**, Schematic representation of the photocycle. The absorption maxima for the ground (bR₅₇₀) and the principal intermediate states (in nm) are shown in subscripts; the respective lifetimes at room temperature, as well as deprotonation and reprotonation, are shown next to the arrows. **b**, The all-*trans* and 13-*cis* configurations of retinal in bacteriorhodopsin. The latter is formed upon photoisomerization. The chromophore is linked covalently to Lys 216 through a Schiff base.

[†] Present address: Department of Biochemistry and Biophysics, UCSF, San Francisco, California 94143-0448, USA

Table 1 Diffraction data

Data set	Ground*	Excited 1†‡	Excited 2‡	Excited 3‡
Resolution (Å)	38–1.9	38–2.1	38–2.3	23–2.2
No. of observations	97807	67437	36312	42107
No. of unique reflections	17996	13653	10351	11670
Completeness (%)	99.5	99.8	97.2	99.0
(outer shell)§	(98.3)	(99.6)	(95.0)	(99.3)
R_{sym} (%)	4.6	5.5	6.6	5.5
(outer shell)	(28.7)	(37.0)	(37.8)	(27.9)
Mean $I/\sigma(I)$	9.9	10.1	8.0	9.1
(outer shell)	(2.6)	(2.1)	(1.7)	(2.4)
Unit cell (Å)¶	60.80, 60.80, 110.52	61.06, 61.06, 110.57	60.82, 60.82, 110.80	60.75, 60.75, 110.87
Twining (%)	<2	<2	<2	<2

* Identical to the observations from which PDB entry 1qjh was refined¹⁵.

† This data set was used in the refinement.

‡ These data sets were collected from three different crystals.

§ Outer shells were 2.0/1.9, 2.2/2.1, 2.4/2.3 and 2.3/2.2 Å for the ground state and the three excited states, respectively. R_{merge} on the intensities between the ground state and the three excited states was 9.8, 10.1 and 8.3%, respectively.

|| $R_{\text{sym}} = \sum_i \sum_h |I(h) - \langle I(h) \rangle| / \sum_i \sum_h I(h)$.

¶ The space group is $P6_3$.

Table 2 Refinement statistics

	Alternate conformation refinement*	Selected-atom refinement
Data set	Excited 1	Extrapolated
Refinement	Ground + K_{LT}	K_{LT}
Refined residues	228	10
Residues (water molecules) with dual conformations	10† (2)	0
Total no. of water molecules	26	25
R_{cryst} (%)‡	22.29	26.07
R_{free} (%)§	25.48	30.33
R.m.s.d. on		
bond lengths (Å)	0.0075	0.0137
bond angles (°)	1.14	1.25
Average B -factor ($\text{\AA}^2$)	35.8	39.0
K_{LT} occupancy (%)	35	100

* To ascertain that the structure of the ground state population in the photoexcited crystal remains unchanged, the refinement was also performed without fixing the ground-state contribution.

† The refinement was also extended to 28 dual conformations: the 18 additional residues did not move significantly.

‡ $R_{\text{cryst}} = \sum_i |F_o(h)| - |F_{\text{calc}}(h)| / \sum_i |F_o(h)|$.

§ || R.m.s.d., root-mean-square deviation.

the ground state (Fig. 2a) and the photoexcited intermediate (Fig. 2b) trapped in crystals of bacteriorhodopsin. When bacteriorhodopsin crystals were illuminated with green light ($\lambda = 532$ nm) at 110 K, the difference absorption spectrum shown in Fig. 2b, representing the low-temperature K intermediate⁸ (K_{LT}), was produced, in accordance with the build up of K_{LT} within purple membranes under similar conditions^{9,10}. Similar difference spectra were recorded in the temperature range of 110 to 125 K. The correspondence between the K_{LT} intermediate and K (the intermediate at physiological temperature) is established by spectroscopy, as both K and K_{LT} show a red shift in the absorption maximum as compared with the ground state spectrum. Resonance Raman spectroscopy has shown that, at room temperature, the K intermediate adopts a planar 13-*cis* retinal configuration⁸, whereas the 13-*cis* retinal geometry is slightly distorted for K_{LT} ^{8,11}. Raising the temperature of the crystal above 125 K leads to depletion of the K_{LT} state, followed by formation of the L state and eventually completion of the photocycle of bacteriorhodopsin at much higher temperatures. We collected monochromatic X-ray diffraction data

sets on a number of ground states (F_{ground}) and photoexcited states (F_{exc}) at 110 K, with the crystals being continuously illuminated during data collection in the latter cases (see Methods). The main conclusions were drawn from $(F_{\text{exc}} - F_{\text{ground}})$ difference Fourier maps and were confirmed by refinement.

Light-induced structural changes are confined to the vicinity of the chromophore (Fig. 3), consistent with our spectral observation that the photocycle did not progress beyond K_{LT} . The small magnitude of these movements could not be observed by electron diffraction at 3.5 Å resolution in projection maps¹⁰. Along the retinal, the strongest difference appears as a negative electron density tangential to the C14–C15 bond (Fig. 4a). The absence of an equally strong positive feature indicates increased disorder in the vicinity of the isomerized C13=C14 bond, presumably due to thermal effects (532-nm photons deposit 225 kJ mol⁻¹ of energy into the retinal). However, the bulk of the retinal remains locked in place, consistent with earlier spectral observations¹². Photoexcitation of the retinal induces movements in the side chain of Lys 216, seen in Fig. 4a and 4b as complementary positive and negative densities aligned parallel to this side chain's ground-state orientation. This rearrangement is accompanied by a movement of the peptide backbone of Lys 216 in helix G (Fig. 4b), disrupting the hydrogen bond from Lys 216 to Gly 220, in agreement with the observation that the carbonyl group of Lys 216 becomes solvated during the photocycle¹³. Changes in the orientation of the Schiff base dislocate a key water molecule (W402, Fig. 4c) which, in the ground-state structures^{14,15}, is situated between the positively charged Schiff base and the negatively charged carboxylate of Asp 85. The breaking of hydrogen-bonding interactions between this water molecule and the Oδ1 atom of Asp 85 increases the pK_a of this residue¹⁶ and enables Oδ1 of Asp 85 to move closer to the position originally occupied by the Schiff base (Fig. 4c). Associated with this rearrangement is a translation of the peptide backbone of Asp 85 in helix C. A second water molecule (W401, Fig. 4c), which also forms a hydrogen bond to Asp 85 in the ground-state structure, is partially disordered, and Tyr 57 moves towards the extracellular side, maintaining the hydrogen bond to Asp 212.

From the difference electron-density maps calculated between the ground state and the three excited-state data sets (Table 1), Met 20, Tyr 57, Asp 85, Asp 212, Lys 216, Ret 300 (the retinal), W401 and W402 were identified as displaying the most pronounced photo-induced movements. The ground-state bacteriorhodopsin structure¹⁵ (protein data bank (PDB) entry 1qjh with the lipids removed) was subjected to a rigid body refinement using F_{exc} . Simulated annealing was then performed with the above residues and adjacent residues (Ala 84, Trp 86, Ala 215 and Val 217) held fixed to their ground-state coordinates. Subsequent refinement of

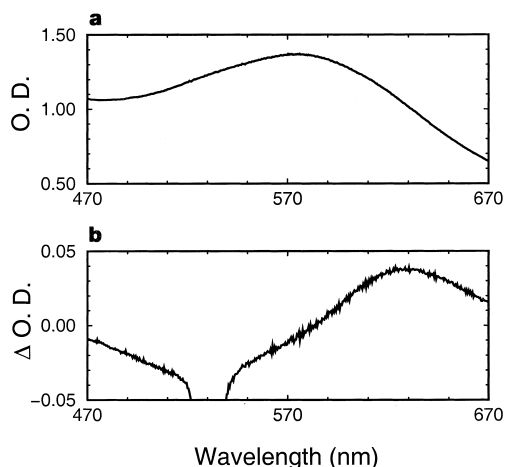


Figure 2 Spectral characterization of a single crystal of bacteriorhodopsin. **a**, Absorption spectrum of a typical light-adapted bacteriorhodopsin crystal in the ground state at 110 K. **b**, Difference spectrum between the spectrum of the crystal when photoexcited with green light ($\lambda = 532$ nm) and the spectrum shown in **a**. The spectrum in **b** is characteristic of the low-temperature K intermediate (K_{LT})^{9,10}. A discontinuity at 532 nm arises from the scatter of laser light. From bacteriorhodopsin spectra⁹ recorded at 110 K, the population of the K_{LT} state within the crystal at this illumination intensity was estimated to be 20% after baseline corrections. Other crystals were used for X-ray data collection. O.D., optical density.

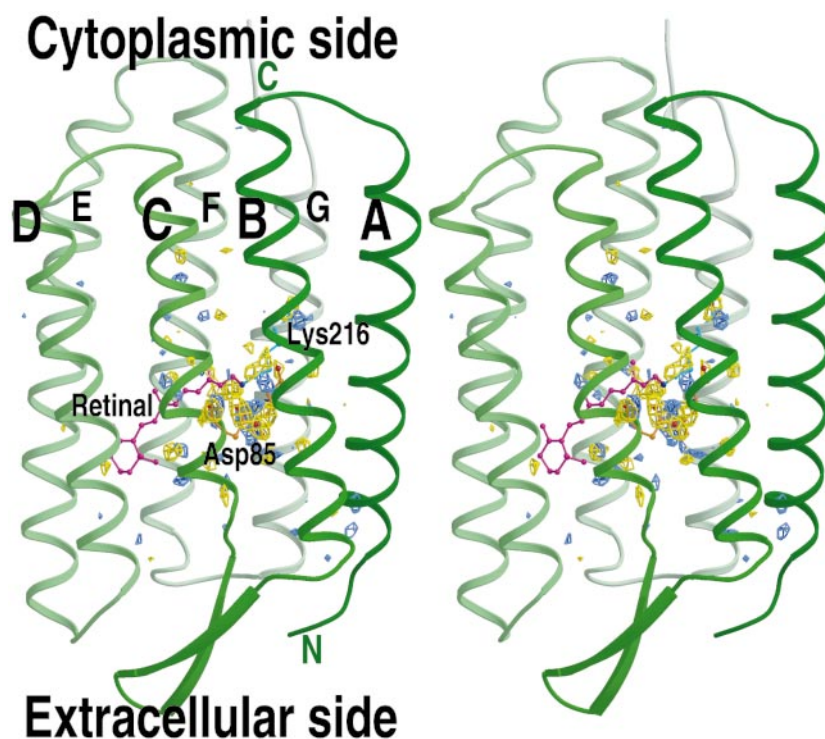


Figure 3 Structural changes in bacteriorhodopsin resulting from photoexcitation at 110 K. Stereo overview of the ($F_{\text{exc}} - F_{\text{ground}}$) difference electron-density maps (yellow, negative; blue, positive density) contoured to 3.5σ (σ , root mean square electron density for the unit

cell) and overlaid on the ground-state bacteriorhodopsin model. At this contour level most changes are clustered in the vicinity of the retinal.

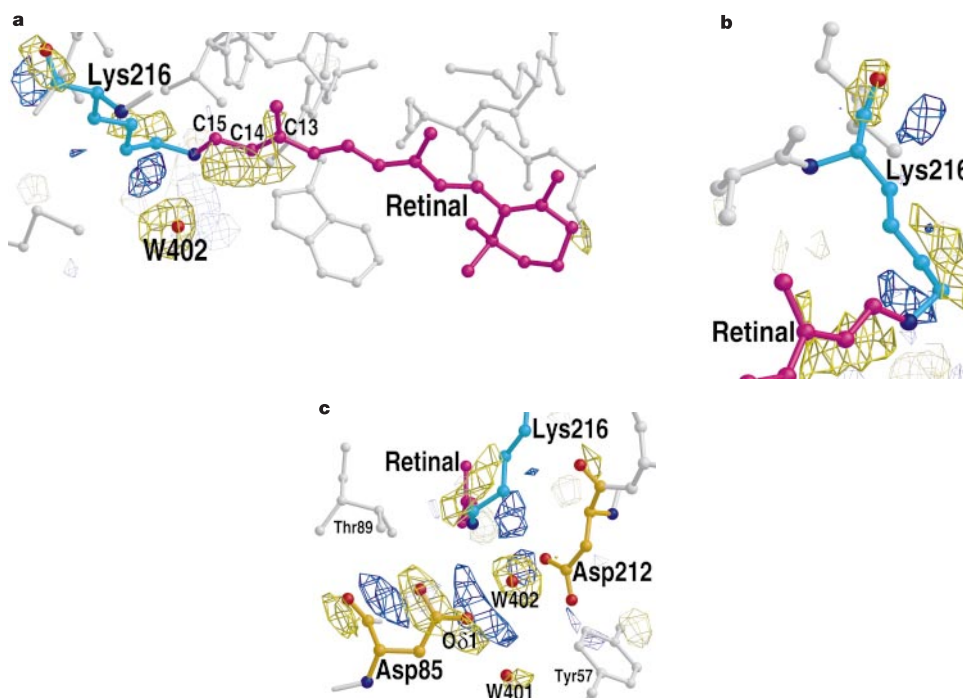


Figure 4 Detailed view of the observed electron-density differences (yellow, negative; blue, positive) overlaid on the ground-state structure, contoured to 3.5σ . **a**, Retinal (purple) showing changes around C14–C15 and Lys 216 (cyan), and the negative peak at the position of water W402. **b**, View indicating the movement of the peptide backbone of Lys 216 and movements along its side chain. **c**, Movements of Asp 85 (orange), Asp 212

(orange) and the water molecule W402 near the Schiff base. In the ground-state structure W402 forms hydrogen bonds to the O δ 1 atom of Asp85 and the nitrogen of the Schiff base (N ϵ of Lys 216, dark blue). Partial disordering of W401 and movement of Tyr 57 are also visible.

the atomic positions of all other residues, and individual B -factor refinement on all atoms, yielded R_{cryst} and R_{free} values of 22.32 and 25.54%, respectively. A model for the low-temperature intermediate was then built into the difference electron-density map ($F_{\text{exc}} - F_{\text{ground}}$) and was added as an alternative conformer with an occupancy of 35% (see Methods). Further refinement of the atomic positions of the selected residues of the alternative conformer, and subsequent B -factor refinement of all atoms, resulted in an R_{cryst} of 22.29% and an R_{free} of 25.48% (Table 2; Fig. 5). We also carried out selected-atom refinement against extrapolated structure factor amplitudes¹⁷ for the low-temperature intermediate: $F_{\text{extrapolated}} = (F_{\text{exc}} - 0.65 F_{\text{ground}})/0.35$ (Table 2; Fig. 5). Both refinement approaches yielded similar structural results (root-mean-square deviations of 0.22 Å on the refined atoms). The calculated electron-density difference ($F_{\text{exc}} - F_{\text{ground}}_{\text{calc}}$) displays the same features as the experimental difference map.

On the basis of these results, we propose the following model for the early events in the photocycle of bacteriorhodopsin: photoisomerization of the retinal to the 13-*cis* 15-*anti* configuration (Fig. 5) forces C20 towards the cytoplasmic side of the membrane, thereby bringing it into closer contact with the indole ring of Trp 182. This action has been suggested as a mechanism through which retinal isomerization could induce large movements at the cytoplasmic end of helix F²¹. Changes in the position of Lys 216 and its peptide backbone in helix G also occur. These rearrangements partially disrupt a hydrogen-bonding network and thereby weaken this helix, allowing it to flex at later stages in the photocycle. Significant movements of helices F and G on the millisecond time-scale^{18–22} have been implicated in the mechanism of reprotonation of the Schiff base. Movement of the Schiff base also breaks hydrogen bonds to a water molecule (W402, Fig. 4c) held in the ground state^{14,15} between the nitrogen of the Schiff base and the carboxylates

of Asp 85 and Asp 212. Displacement of this water is accompanied by the mutual approach of the two carboxylates (Fig. 4c), creating an electrostatic environment for effective proton transfer from the Schiff base to Asp 85. The early rearrangements described here initiate a sequence of structural changes spanning almost nine orders of magnitude in time, and resulting in the vectorial translocation of a proton from one side of the purple membrane to the other. □

Methods

Protein crystallization

Detergent-solubilized bacteriorhodopsin was crystallized within the lipidic cubic phase as described^{3,4}. The crystallization matrix was liquefied with lipase²³ before crystal mounting. Crystals of bacteriorhodopsin formed hexagonal plates around 70 µm across and 5–10 µm thick.

Single-crystal microspectrophotometry

Spectra were recorded at various temperatures (92–280 K) from a measurement area of 25 µm diameter, using a portable microspectrophotometer⁷ (<http://www.4dx.se>). The optical density of the crystals was 0.7–1.5 at 570 nm and 0.5–1.2 at 532 nm (the pump wavelength).

Crystal illumination

Bacteriorhodopsin was light adapted in the crystalline state by exposure to bright white light for several minutes before the crystals were mounted in cryo-loops. Crystals were then flash frozen in liquid nitrogen. During X-ray diffraction experiments, crystals at 110 K were continuously excited under similar conditions to those established by off-line spectral characterization (Fig. 2). Light from a 100 mW laser diode ($\lambda = 532$ nm) was guided to the crystal through an optical fibre (225 µm diameter) which passed through the centre of a cryo-top hat (4DX-Rays Systems AB, Uppsala) and pin on which the cryo-loop was fixed. The loop was bent so as to align the crystal at an angle of $\sim 60^\circ$ relative to the fibre tip, with an estimated light flux of 30 W cm⁻² at the position of the crystal. This system was rotated with the crystal during data collection, maintaining a constant illumination geometry throughout the experiment.

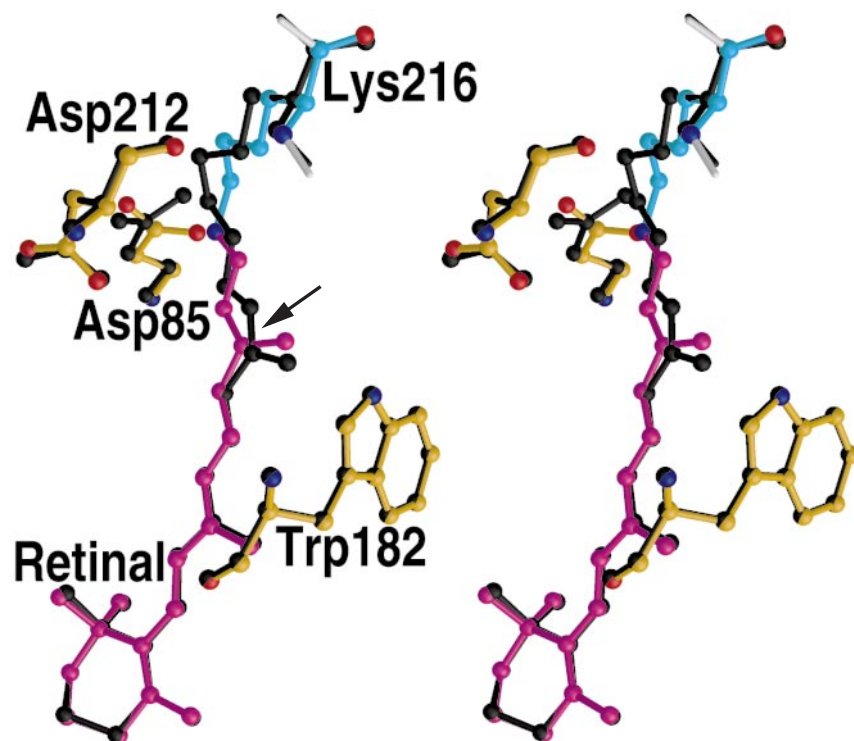


Figure 5 Stereo view of the structure of the early intermediate in the bacteriorhodopsin photocycle. The model was obtained by refinement against extrapolated data (see Methods). The refined atomic positions in K_{L}^+ (black) near the retinal (Table 2) are

superimposed on the ground-state structure (same colours as Fig. 4; W402 is omitted). The bond undergoing photoisomerization (C13=C14) is highlighted by an arrow.

X-ray data collection

X-ray diffraction data ($\lambda = 0.94 \text{ \AA}$) were collected at beamline ID14-EH3 of the European Synchrotron Radiation Facility (ESRF), Grenoble, France, using a MAR CCD X-ray detector (MARResearch, USA). The life time of crystals within the X-ray beam was sufficient to allow the collection of one complete data set per crystal, consisting of 90 frames of 1° oscillation and 30 s exposure per frame. Heating by the X-ray beam during data collection was estimated to be less than 1 K. We collected data sets from a number of excited state and ground state crystals.

Data processing and structural analysis

Data were processed using the HKL package²⁴ and the CCP4 suite²⁵. The degree of merohedral twinning was determined^{15,26,27} for each data set using Britton plots and Yeates statistics, and only data sets with twinning ratios $<2\%$ were used (statistics for these data sets are shown in Table 1). Phases from the ground state¹⁵ (PDB entry 1qhj) were used to calculate ($F_{\text{exc}} - F_{\text{ground}}$) electron-density difference maps. We optimized signal-to-noise ratios for Fourier difference maps using a Bayesian statistics-based method²⁸. Electron-density differences were calculated between the three excited and the ground-state data sets and the features described here were found to be reproducible. As electron-density difference maps display considerably less bias than a refined structure, they were used extensively to interpret the results, and the structural rearrangements described above were corroborated by structure refinement using CNS²⁹. Initial models for K_{LT} were built interactively using O³⁰. Figures were drawn using a modified version of MOLSCRIPT³¹ and were rendered using Raster3D³².

Crystallographic determination of the occupancy of K_{LT}

We used the ground-state structure as a reference from which to establish the occupancy of the K_{LT} intermediate. By setting the B -factor of W402 to its ground-state value¹⁵ and refining this water's occupancy against F_{exc} , we estimated a minimum population of the ground state within the illuminated crystal to be 60%. Simulated annealing omit maps (with the active-site residues associated with peaks in the electron-density differences omitted) were calculated from extrapolated structure factor amplitudes ($F_{\text{exc}} - \alpha F_{\text{ground}})/(1 - \alpha)$ for a range of α . Selected-atom refinement of the K_{LT} structure against $F_{\text{extrapolated}}$ was then performed for $\alpha = 0.60, 0.65, 0.70$ and 0.75 , with various weightings of the X-ray observations relative to the stereochemical terms¹⁷. A combination of R_{cryst} , R_{free} , the average B -factor and the quality of $2F_{\text{exc}} - F_{\text{calc}}$ electron density maps gave an optimum K_{LT} population of around 35%, consistent with population estimates from off-line microspectrophotometry measurements.

Retinal constraints

At $2.1 \text{ \AA}$ resolution, the two closely overlapping retinal configurations (all-*trans*, 15-*anti* and 13-*cis*, 15-*anti*) could not be refined free from bias resulting from stereochemical constraints. The retinal backbone and Schiff base linkage up to C ϵ of Lys 216 were constrained to be planar for both the all-*trans*, 15-*anti* and the 13-*cis*, 15-*anti* configurations.

Received 20 April; accepted 9 September 1999.

- Oesterhelt, D. & Stoekenius, W. Rhodopsin-like protein from the purple membrane of *Halobacterium halobium*. *Nature New Biol.* **233**, 149–152 (1971).
- Henderson, R. & Unwin, P. N. T. Three dimensional model of purple membrane obtained by electron microscopy. *Nature* **257**, 28–32 (1975).
- Landau, E. M. & Rosenbusch, J. P. Lipidic cubic phases: A novel concept for the crystallization of membrane proteins. *Proc. Natl Acad. Sci. USA* **93**, 14532–14535 (1996).
- Rummel, G. *et al.* Lipidic cubic phases: New matrices for the three dimensional crystallization of membrane proteins. *J. Struct. Biol.* **121**, 82–91 (1998).
- Schlichting, I., Berendzen, J., Phillips, G. N. Jr & Sweet, R. M. Crystal structure of photolysed carbonmonoxy-myoglobin. *Nature* **371**, 808–812 (1994).
- Genick, U. K., Soltis, S. M., Kuhn, P., Canestrelli, I. L. & Getzoff, E. D. Structure at $0.85 \text{ \AA}$ resolution of an early protein photocycle intermediate. *Nature* **392**, 206–209 (1998).
- Hadfield, A. & Hajdu, J. A fast and portable microspectrophotometer for protein crystallography. *J. Appl. Crystallogr.* **26**, 839–842 (1993).
- Doig, S. J., Reid, P. J. & Mathies, R. A. Picosecond time-resolved resonance Raman spectroscopy of bacteriorhodopsin's J, K and KL intermediates. *J. Phys. Chem.* **95**, 6372–6379 (1991).

- Xie, A. Quantum efficiencies of bacteriorhodopsin photochemical reactions. *Biophys. J.* **58**, 1127–1132 (1990).
- Bullough, P. A. & Henderson, R. The projection structure of the low temperature K intermediate of the bacteriorhodopsin photocycle determined by electron diffraction. *J. Mol. Biol.* **286**, 1663–1671 (1999).
- Braiman, M. & Mathies, R. Resonance Raman spectra of bacteriorhodopsin's primary photoproduct: Evidence for a distorted 13-*cis* retinal chromophore. *Proc. Natl Acad. Sci. USA* **79**, 403–407 (1982).
- Song, Q., Harms, G. S., Wan, C. & Johnson, C. K. Reorientations in the bacteriorhodopsin photocycle. *Biochemistry* **33**, 14026–14033 (1994).
- Takei, H., Gat, Y., Rothman, Z., Lewis, A. & Sheves, M. Active site lysine backbone undergoes conformational changes in the bacteriorhodopsin photocycle. *J. Biol. Chem.* **269**, 7387–7389 (1994).
- Luecke, H., Richter, H. T. & Lanyi, J. K. Proton transfer pathways in bacteriorhodopsin at $2.3 \text{ \AA}$ resolution. *Science* **280**, 1934–1937 (1998).
- Belrhali, H. *et al.* Protein, lipid and water organization in bacteriorhodopsin crystals: a molecular view of the purple membrane at $1.9 \text{ \AA}$ resolution. *Structure* **7**, 909–917 (1999).
- Gat, Y. & Sheves, M. A mechanism for controlling the pKa of the retinal protonated Schiff base in retinal proteins. A study with model compounds. *J. Am. Chem. Soc.* **115**, 3772–3773 (1993).
- Genick, U. K. *et al.* Structure of a protein photocycle intermediate by millisecond time-resolved crystallography. *Science* **275**, 1471–1475 (1997).
- Dencher, N. A., Dresselhaus, D., Zaccari, G. & Büldt, G. Structural changes in bacteriorhodopsin during proton translocation revealed by neutron diffraction. *Proc. Natl Acad. Sci. USA* **86**, 7876–7879 (1989).
- Koch, M. H. J. *et al.* Time-resolved X-ray diffraction study of structural changes associated with the photocycle of bacteriorhodopsin. *EMBO J.* **10**, 521–526 (1991).
- Nakasako, M., Kataoka, M., Amemiva, Y. & Tokunaga, F. Crystallographic characterization by X-ray diffraction of the M-intermediate from the photocycle of bacteriorhodopsin at room temperature. *FEBS Lett.* **292**, 73–75 (1991).
- Subramaniam, S., Gerstein, M., Oesterhelt, D. & Henderson, R. Electron diffraction analysis of structural changes in the photocycle of bacteriorhodopsin. *EMBO J.* **12**, 1–8 (1993).
- Vonck, J. A three-dimensional difference map of the N-intermediate in the bacteriorhodopsin photocycle: part of the F helix tilts in the M to N transition. *Biochemistry* **35**, 5870–5878 (1996).
- Nollert, P. & Landau, E. M. Enzymic release of crystals from lipidic cubic phases. *Biochem. Soc. Trans.* **26**, 709–713 (1998).
- Otwiński, Z. in *Data Collection and Processing* (eds Sawyer, L., Isaacs, N. W. & Bailey, S.) DL/SCI/R34, 55–62 (Daresbury Laboratory, Warrington, 1993).
- Collaborative Computational Project nr4. The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763 (1994).
- Yeates, T. O. Detecting and overcoming crystal twinning. *Methods Enzymol.* **276**, 344–358 (1997).
- Valegård, K. *et al.* Structure of a cephalosporin synthase. *Nature* **394**, 805–809 (1998).
- Ursby, T. & Bourgeois, D. Improved estimation of structure-factor difference amplitudes from poorly accurate data. *Acta Crystallogr. A* **53**, 564–575 (1997).
- Brünger, A. T. *et al.* Crystallography and NMR system: A new software system for macromolecular structure determination. *Acta Crystallogr. D* **54**, 905–921 (1998).
- Jones, T. A., Bergdoll, M. & Kjeldgaard, M. in *Crystallographic and Modelling Methods in Molecular Design* (eds Bugg, C. & Ealick, S.) 189–199 (Springer, New York, 1990).
- Esnouf, R. M. An extensively modified version of MolScript that includes greatly enhanced coloring capabilities. *J. Mol. Graph.* **15**, 133–138 (1997).
- Merritt, E. A. & Murphy, M. E. P. Raster3D version 2.0. A program for photorealistic molecular graphics. *Acta Crystallogr. D* **50**, 869–873 (1994).

Acknowledgements

We thank D. Bourgeois, L.-O. Essen, R. Henderson, D. Oesterhelt, J. P. Rosenbusch, I. Schlichting and S. Subramaniam for discussions; W. P. Burnmeister, A. Hardmeyer, T. Taylor and R. Wouts for experimental contributions; and G. Büldt for providing purple membrane. Support from the EU-BIOTECH, the Swedish Research Council NFR, the Institut Universitaire de France and the Swiss National Science Foundation's SPP BIOTECH is acknowledged. K.E. was an undergraduate student in the Molecular Biotechnology Programme of Uppsala University.

Correspondence and requests for materials should be addressed to E.M.L. (e-mail: landau@ubaclu.unibas.ch), R.N. (e-mail: neutze@xray.bmc.uu.se) or E.P.P. (e-mail: pebay@ibs.fr). Crystallographic coordinates and structure factor amplitudes for all data sets have been deposited with the Protein Data Bank (accession codes: 1qko and 1qkp).

Touching a nerve

Antibodies feature strongly in a selection aimed at neuroscientists.



Matters of substance P: antagonists galore.

Neurotransmissions newsletter

Sigma RBI www.sigma-aldrich.com

Latest edition of this free quarterly

Neurotransmissions Vol. 15, No. 2, published by Sigma RBI, includes an article entitled 'Use of Substance P Receptor Antagonists as Research Tools in Psychopharmacology', by Nadia Rupniak. New Sigma-RBI products featured include the serotonin 5-HT₆ antagonist, Ro 04-6790 dihydrochloride, and the 5HT_{1B} antagonist, SB 224289 hydrochloride.

Reader Service No. 100

Catecholamine extraction

Jones Chromatography www.joneschrom.com

An application note on the extraction of catecholamines

This note covers catecholamine extraction with the 4- μ m Genesis C18 HPLC column. Compounds extracted include noradrenaline, adrenaline, L-dopa, dopamine, and isovanillyl mandelic acid. Columns come in C18, C8, C4, CN, amino, silica and phenyl phases.

Reader Service No. 101

Long Evans rat

Taconic www.taconic.com

Born-again, and again, rats

Taconic introduced this rat earlier this year. It is an outbred, hooded rat with pigmented eyes, used for neurological, toxicological and ophthalmologic studies. The line was estab-

lished by Long and Evans in 1915 by crossing several Wistar Institute white females with a wild grey male. Taconic's model was rederived in 1975 by Simonsen Labs, using stock obtained from the University of California at Berkeley in 1949. Taconic again rederived the type in August 1998 and it now meets the company's Murine Pathogen Free standard.

Reader Service No. 102

Neuropeptide Y

DiaSorin

www.diasorin.com

Affinity-purified rabbit polyclonal antibody to Neuropeptide Y Y1 receptor

Neuropeptide Y (NPY) neurons are important mediators of a variety of biological activities in the brain and spinal cord, the peripheral nervous system, and the gastrointestinal tract. The Y1 receptor mediates the effects of this peptide neurotransmitter and is one of the most important and widespread. The DiaSorin antiserum to NPY Y1 receptor was generated in rabbit against the C terminus of the rat Y1 receptor.

Reader Service No. 103

KOR-1 antibody

Upstate Biotechnology

www.upstatebiotech.com

A polyclonal antibody to KOR-1

Kappa Opioid Receptor (KOR-1) receptors have been shown to be one of the receptors involved in binding products resulting from morphine metabolism. The immunogen used to generate this antibody corresponds to amino acids 300–312 of the extracellular loop of KOR-1.

Reader Service No. 104

Microelectrode holder

ALA Scientific Instruments

www.alascience.com

Get a grip on glass microelectrodes

As the link between the living cell and high-impedance amplifiers, microelectrode hold-



Glass act: microelectrode holders to order.

ers are vital components in many experimental set-ups. Applications for glass microelectrodes include: intracellular and extracellular recordings, whole-cell and single-cell patch-clamp recordings, iontophoresis and ion-specific measurements. ALA Scientific produce a wide range of holders, and can fabricate designs to meet individual requirements

Reader Service No. 105

Clone 013

Zymed Laboratories

www.zymed.com

For studying Ewing's sarcoma

This new monoclonal antibody is useful for detecting the MIC gene product (p30/32^{MIC2}) which is reported to be consistently overexpressed in Ewing's sarcoma and peripheral neuroectodermal tumours. The MIC gene product is a cell-surface glycoprotein and is thought to be involved in cell to cell interactions. Clone 013 is useful as part of an antibody panel for distinguishing Ewing's sarcoma and peripheral neuroectodermal tumours from lymphomas.

Reader Service No. 106

AChR antibodies

Covance

www.babco.com

Receptor antibodies, for work on nicotine, Alzheimer's, Parkinson's, myasthenia gravis

Nicotinic acetylcholine receptors (AChRs) are ligand-gated ion channels located at

ADVERTISEMENT



Shaping the future of spectrophotometry

- **lightens workloads**
- **increases throughput**
- **creates more space**

www.wpald.co.uk

lightwave
UV/Vis Diode Array Spectrophotometer

WPA



neuromuscular junctions or neuron-neuron synapses. Covance have introduced 16 new monoclonal antibodies against individual AChR subunits, including muscular and neuronal forms. These antibodies have broad species specificity and are applicable for immunoblotting, immunoprecipitation and immunofluorescence.

Reader Service No. 107

Anti-BIII Tubulin

Promega www.promega.com

A marker for neurons

Anti-BIII Tubulin monoclonal antibody has been added to Promega's Neuroscience product line. It is a Protein G purified IgG, monoclonal from the clone 5G8. It is directed against the peptide region (EAQGPK) corresponding to the C terminus of BIII tubulin, a specific marker for neurons. The antibody is recommended for labelling neurons in tissue sections and cell culture, western blotting and immunostaining.

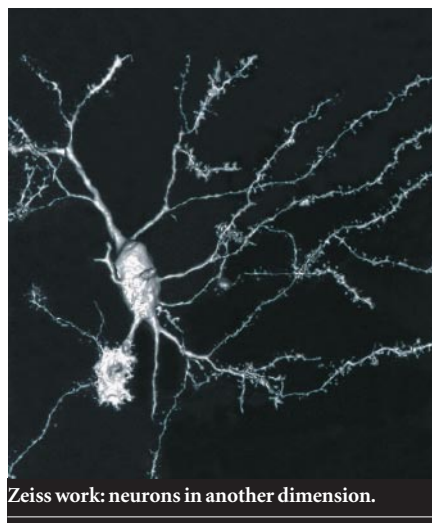
Reader Service No. 108

3DforLSM

Carl Zeiss www.zeiss.de

Software for evaluation of LSM image stacks

3DforLSM is now available for users of Zeiss laser scanning microscopes. It allows the 3D visualization and measurement of structures and objects, achieving genuine 3D quantitative determination of absolute values, for



Zeiss work: neurons in another dimension.

example of the volume, the surface and average gray value. Comparative analyses can also be performed in the third dimension. The software runs two 3-D reconstruction techniques, surface rendering and alpha rendering.

Reader Service No. 109

Ion channel antibodies

Chemicon www.chemicon.com

A range of antibodies to ion channels

Ion channels for which Chemicon can provide antibodies include calcium, chloride, potassium and sodium. Also in the catalogue are antibodies against P2X receptors, PDZ domain proteins and muscarinic receptors.

Each of these polyclonal antibodies includes a free vial of control antigen.

Reader Service No. 110

Neuroscience antibodies

Biogenesis www.biogenesis.co.uk

Antibodies for research in the fields of Alzheimer's and Parkinson's diseases

This new range includes polyclonal antibody to a synthetic peptide corresponding to amino acids 116-131 of the C terminus of the Parkinson's disease α -synuclein protein. The α -synuclein gene, on chromosome 4, is said to be responsible for early-onset autosomal dominant Parkinson's disease.

Reader Service No. 111

Rat GEM 2

Incyte Pharmaceuticals www.incyte.com

The company's second rat-specific GEM gene expression microarray

Each high-density Rat GEM 2 microarray contains over 8,400 genes known to be active in the rat central nervous system, as well as genes believed to be important in rat neuronal functions. There are genes from 18 different rat nervous tissues on this microarray, including cerebrum, whole brain and spinal cord.

Reader Service No. 112

These notes are compiled in the Nature office from information provided by the manufacturers. For more details, fill in the reader service card.

ADVERTISEMENTS

Customised monoclonal & polyclonal antibodies with guarantee!

BioGenes GmbH
Köpenicker Str. 325 • D-12555 Berlin
Tel. +49-30-65 76 23 96 Fax: - 97
E-mail: BioGenes@Berlin.sireco.net
Internet: <http://www.biogenes.de>

READER ENQUIRY NO. 17

Homogenizers?

- Tissue-Tearor™
- BeadBeater™
- Mini-BeadBeater™
- BioHomogenizer™

Check Us Out

www.biospec.com

BioSpec Products, Inc.
Ph/Fax 918-336-3363
e-mail info@biospec.com

New Probe Sizes
0.45cm

READER ENQUIRY NO. 54

Detect residual DNA with Wako's DNA Extractor Kit or Isolate Genomic DNA with DNA Extractor WB Kit

Both use the Nal Method
No Phenol No Chloroform

Wako

Wako BioProducts
Wako Chemicals USA, Inc.
1600 Bellwood Road
Richmond, VA 23237
Telephone: (804) 271-7677
Facsimile: (804) 271-7791
(800) 992-9256
bioproducts@wakousa.com

READER ENQUIRY NO. 39

CUSTOM PEPTIDES

HPLC MS MALDI-TOF

- Up to 30 aa
- HPLC & MS data

Purity	>80%	>95%	>99%
Amount	15mg	10mg	

info@genosphere-biotech.com
Tel: +33 1 42 71 70 21 • Fax: +33 1 42 71 70 23

GENOSPHERE BIOTECHNOLOGIES

READER ENQUIRY NO. 24

Diagnostic Reagents

Pefachrome®-Series: chromogenic and fluorogenic substrates

Pefabloc®-Series: serine proteinase inhibitors

Purified snake venom proteins (e.g. Protac®, Ecarin, Batroxobin)

PENTAPHARM LTD
CH 4002 Basel Fax: +41 61 319 96 19
<http://www.pentapharm.ch>

READER ENQUIRY NO. 6

Long Read DNA Sequencing

Up To 1200 bp/read

Fast Turnaround Time

Only \$50/read

MacConnell Research Corp
1-800-466-7949

READER ENQUIRY NO. 27